Technical Report 1077

# Predicting Performance Ratings Using Motivational Antecedents

**Michelle M. Zazanis and Martha S. Lappin**
U.S. Army Research Institute



**January 1998**


**United States Army Research Institute for the Behavioral and Social Sciences**

**U.S. Army Research Institute**
**for the Behavioral and Social Sciences**

**A Directorate of the U.S. Total Army Personnel Command**

**EDGAR M. JOHNSON**
**Director**

---

Technical Review by

Robert N. Kilcullen, ARI
M. A. Fischl, ARI

## NOTICES

**DISTRIBUTION:** Primary distribution of this Technical Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: TAPC-ARI-PO, 5001 Eisenhower Ave., Alexandria, VA 22333-5600.

**FINAL DISPOSITION:** This Technical Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research for the Behavioral and Social Sciences.

**NOTE:** The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

# REPORT DOCUMENTATION PAGE

| 1. REPORT DATE (dd-mm-yy) | 2. REPORT TYPE | 3. DATES COVERED (from. . . to) |
|---|---|---|
| January 1998 | Final | March 1995 – December 1997 |

| 4. TITLE AND SUBTITLE | 5a. CONTRACT OR GRANT NUMBER |
|---|---|
| Predicting Performance Ratings Using Motivational Antecedents | 5b. PROGRAM ELEMENT NUMBER<br>63007A |

| 6. AUTHOR(S) | 5c. PROJECT NUMBER |
|---|---|
| Michele M. Zazanis and Martha S. Lappin | 2O263007A792 |
| | 5d. TASK NUMBER<br>501 |
| | 5e. WORK UNIT NUMBER<br>H01 |

| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) | 8. PERFORMING ORGANIZATION REPORT NUMBER |
|---|---|
| U.S. Army Research Institute for the Behavioral and Social Sciences<br>ATTN: TAPC-ARI-RP<br>5001 Eisenhower Avenue<br>Alexandria, VA 22333-5600 | |

| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | 10. MONITOR ACRONYM |
|---|---|
| U.S. Army Research Institute for the Behavioral and Social Sciences<br>5001 Eisenhower Avenue<br>Alexandria, VA 22333-5600 | ARI |
| | 11. MONITOR REPORT NUMBER<br>Technical Report 1077 |

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*:

This research examined the role of motivation in predicting peer and trainer ratings of student performance and contrasted the relative importance of various antecedents for peer and trainer ratings. Ability, experience, and self-report personality and belief measures were collected from 239 enlisted male Army soldiers attending training for Special Forces. At the conclusion of the first phase of the training, performance ratings were obtained from peers and trainers. LISREL8 was used to test a path model predicting performance ratings. Results showed observer ratings of effort and self-reported task self-efficacy played a role in predicting ratings of task-specific performance. Self-report measures of mastery beliefs, achievement orientation, and locus of control were not significant contributors to the performance rating models, while the soldier's prior experience had both direct effects on performance as well as indirect effects through self-efficacy and effort. The lack of importance of the personality measures is discussed with respect to the high level of variation among the students on level of prior experience. Finally, analyses indicated that peer raters included more information about interpersonal skills in their rating of overall performance than did trainers. Utilization of these findings with regard to performance evaluation systems is discussed.

15. SUBJECT TERMS Performance ratings; Performance Evaluation; Motivation; Interpersonal skills, Peer evaluations; Peer ratings; Self-efficacy; Special Forces

| SECURITY CLASSIFICATION OF | | | 19. LIMITATION OF ABSTRACT | 20. NUMBER OF PAGES | 21. RESPONSIBLE PERSON (Name and Telephone Number) |
|---|---|---|---|---|---|
| 16. REPORT | 17. ABSTRACT | 18. THIS PAGE | | | |
| Unclassified | Unclassified | Unclassified | Unlimited | 77 | Dr. Michelle Zazanis<br>703/617-0318 |

Technical Report 1077

# Predicting Performance Ratings Using Motivational Antecedents

**Michelle M. Zazanis and Martha S. Lappin**
U.S. Army Research Institute

**Organization and Personnel Resources Research Unit**
**Paul A. Gade, Chief**

**U.S. Army Research Institute for the Behavioral and Social Sciences**
**5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600**

**January 1998**

---

**Army Project Number**
**2O263007A792**

**Manpower and Personnel**

## FOREWORD

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) performs research on personnel performance and training issues of particular significance to the U.S. Army. During FY96, ARI developed new measures for the assessment and selection of soldiers for Special Forces, including new methods for obtaining peer assessments in their assessment and training programs. This report describes an investigation into the personal characteristics that predict the performance ratings a soldier receives from his peers and trainers, focusing on the role of the soldier's motivation and interpersonal skills. The results provide insight into the importance of motivation in performance ratings and into key differences between ratings made by peers and trainers. The results also inform our efforts to both understand and properly utilize the peer evaluation measures we developed for Special Forces selection and training during the course of this research. A description of this research was included in briefings to the Commanding General, U.S. Army John F. Kennedy Special Warfare Center and School, the Department of the Army's Deputy Chief of Staff and Assistant Deputy Chief of Staff for Personnel, the Commander in Chief U.S. Special Operations Command, and the Assistant Secretary of the Army for Manpower and Reserve Affairs.

ARI's participation in this effort is part of an ongoing program of research designed to enhance the quality of Army personnel. The peer evaluation research, in particular, demonstrates the value of collaborative researcher-sponsor efforts to identify, understand, and solve problems existing in the personnel assessment and development process. This work is an essential component of the mission of ARI to conduct research to help manage the force effectively amd efficiently.

ZITA M. SIMUTIS
Technical Director

ACKNOWLEDGMENTS

We would first like to thank several people who were helpful both in providing recommendations during the development of the path models, as well as making suggestions regarding analyses: Stephen Zaccaro, Louis Buffardi, Martin Ford, and Robert Holt. In addition, we would like to thank M.A. Fischl and Robert Kilcullen for their meticulous and insightful reviews. We would also like to thank the Special Forces training company (F Co, 1st Bn, 1st SWTG(A)) for their cooperation and assistance in collecting these data. We hope their new peer evaluation system will be useful to both the cadre and students.

We would like to extend a special note of thanks to Danielle Schweisthal, our diligent research assistant. Her high level of competence in creating and managing databases made this a top quality project, and her high levels of enthusiam and motivation made it an enjoyable experience.

PREDICTING PERFORMANCE RATINGS USING MOTIVATIONAL ANTECEDENTS

## EXECUTIVE SUMMARY


Research Requirement:

To identify the role of motivation in predicting peer and trainer ratings of student performance and contrast the relative importance of various antecedents for peer and trainer ratings; to provide information that can assist the U.S. Special Operations Command in the assessment of soldiers for selection and training in Special Forces.

Procedure:

Ability, experience, and self-report personality and belief measures were collected from 239 enlisted male Army soldiers attending the first phase of training for Special Forces between March and October of 1995. At the conclusion of the first phase of the training, performance ratings were obtained from peers and trainers. A path model was developed to investigate the role of motivational measures, including effort/persistence, self-efficacy, mastery beliefs, achievement orientation, and locus of control, in predicting performance ratings. LISREL8 was used to compare model fit and path coefficients for ratings made by peers and trainers. In addition, hierarchical regression was used to determine the importance of ratings of task-specific performance, interpersonal performance, and effort to overall peer and trainer ratings of current performance and predicted future performance.

Findings:

Results of the path analysis showed observer ratings of effort and self-reported task self-efficacy played a role in predicting ratings of task-specific performance. Self-report measures of mastery beliefs, achievement orientation, and locus of control were not significant contributors to the performance rating models, while the soldier's prior experience had both direct effects on performance as well as indirect effects through self-efficacy and effort. The lack of importance of the self-report personality measures to the models is discussed with respect to the high level of variation that existed in the prior experience of the students.

The comparison of peer and trainer models of performance suggested the antecedents for the performance ratings made by peers were primarily the same as those for trainers. Self-efficacy, however, had a significant direct effect only on peer ratings of task performance. Discussion suggests that situational factors exist that may have attenuated other differences between the peer and trainer models.

Results comparing the importance of task performance, interpersonal performance, and effort to overall performance indicated that peer raters included more information about interpersonal skills in their rating of overall performance than did trainers. In addition, when predicting a soldier's future

performance, only peers significantly increased the importance of interpersonal skills and effort to the ratings.

Utilization of Findings:

Performance ratings are used throughout the Army to evaluate personnel performance; understanding their antecedents and the nature of their dimensionality is, therefore, of critical importance. The rating models in this research have advanced our understanding of the determinants of performance ratings, and have emphasized the importance of motivation to ratings of performance. Specifically, they identified the importance of self-efficacy and effort to performance, and the utility of further examining mastery beliefs. Altering an individual's self-efficacy or mastery beliefs may be a critical step to improving his/her level of effort in a training program. Most importantly, this research has identified for the U.S. Special Operations Command the criticality of the information provided by peer evaluations in Special Forces training. It has suggested that peer evaluations may provide the most critical and irreplaceable information in situations that have a high requirement for interpersonal skills, or that have difficult or ambiguous tasks requiring high levels of effort or persistence.

PREDICTING PERFORMANCE RATINGS USING MOTIVATIONAL ANTECEDENTS

CONTENTS

## LIST OF TABLES



## LIST OF FIGURES

# PREDICTING PERFORMANCE RATINGS USING MOTIVATIONAL ANTECEDENTS

## INTRODUCTION

Understanding the performance rating process is of paramount concern to research in every area of industrial psychology. Ratings are the most commonly used performance criteria, and are used in areas such as evaluating the effectiveness of training, determining promotion decisions, and providing a basis for compensation systems. Often research in these domains focuses on the predictor side of the equation, overlooking the importance of the rating content to the conclusions being drawn. The performance rating process has been examined by researchers in order to further our understanding of the content and structure of ratings.

### Models Predicting Performance Ratings

Early research in this area focused on understanding the relationships among performance ratings and (a) general cognitive ability scores, (b) work sample test performance, and (c) ratee job knowledge test scores (for reviews, see Borman, 1991; Hunter, 1983). A meta-analysis examining the relationships among these variables suggested that ratee job knowledge had the largest direct effect on performance ratings, and that cognitive ability affected ratings through both job knowledge and work samples (Hunter, 1983). In a review of this work, Guion (1983) suggested that these relationships should be expanded to include variables that reflected factors such as ratee characteristics or interpersonal relationship factors.

During the 1980's, the Army's Project A was implemented (See Campbell & Zook, 1991). The goal of this project was to develop a complete system for selecting and classifying Army enlisted personnel, which included the development of a comprehensive set of job performance measures. Three types of performance measures were created: hands-on job sample tests, multiple-choice knowledge tests, and behaviorally anchored rating scales. In addition, performance file data existed that described aspects of an individual's performance in the Army through records of disciplinary actions and awards. The development of this rich database enabled researchers to examine models of the performance rating process using multiple measures of performance as well as multiple performance predictors.

With data from this project, Borman and his colleagues have tested two causal models of performance ratings (Borman, White, Pulakos, & Oppler, 1991; Borman, White, & Dorsey, 1994), and McCloy, Campbell, and Cudeck (1994) have tested a latent model of performance determinants. Borman et al.'s (1991) performance rating model included two ratee characteristics (achievement orientation and dependability), as well as two variables reflecting disciplinary actions and awards. They suggested that, while the cognitive variables that were originally in the Hunter model are maximal ("can-do") measures, the ratee characteristics and disciplinary action and award variables they added reflect more upon typical ("will-do") performance. Will-do performance is affected by motivational and situational factors; factors assumed to be more constant in can-do performance measures.

The addition of these will-do variables resulted in a model that explained more than twice the amount of variance in ratings as that explained by the original Hunter (1983) model (from .16 to .31). This suggests that ratee motivation affected the relationship between ability and performance ratings. Despite the importance of motivational constructs in differentiating between can-do and will-do performance, research has not yet investigated the causal paths leading from ratee motivational variables to performance ratings.

Adding Motivational and Interpersonal Variables

An extensive review of the criterion-related validity of various cognitive and noncognitive predictors demonstrated relationships between performance ratings and motivational constructs such as achievement orientation, dependability, and locus of control (Hough, 1986). Research in the domain of cognitive styles and expectancies has been particularly strong in demonstrating the relationship between belief constructs and motivational behavior as well as performance outcomes. This research, however, on constructs such as helplessness and mastery orientations (Dweck, 1986), and self-efficacy (Bandura, 1986), has predominantly used *objective* performance measures as criteria as opposed to subjective appraisals.

More recent models of performance ratings have added interpersonal variables such as ratee friendliness and "obnoxiousness" (Borman, White, & Dorsey, 1994), and two self-regulatory motivational constructs, goal setting and goal commitment (Barrick, Mount, & Strauss, 1993). Although Barrick, Mount, and Strauss' (1993) inclusion of goal-related constructs does not define a causal path from ratee motivational variables to performance ratings, it does begin to examine the relationship between distal motivational constructs (such as personality measures), proximal motivational constructs (such as goal setting), and performance. As Kanfer (1991) has suggested, it is critical that motivational research employ a unified perspective that specifies the proximal motivational constructs through which more distal personality constructs have their effect on performance.

To summarize, research examining models of performance ratings has offered several critical insights into the nature of these ratings. First, ability and skill factors have a significant impact on ratings of performance, although these factors are accounting for less than 20% of the variance in the ratings. Second, the effect of interpersonal factors must also be considered with respect to ratings. While interpersonal factors are often relevant to job performance, they may be overlooked in delineating performance domains in a job analysis, and would generally not be captured by tests of job knowledge or work sample measures. Finally, motivational aspects of the ratee's performance appear to be critical to performance ratings. It is not clear, however, which ratee motivational characteristics are the strongest determinants of ratings, or which paths mark the effect of these characteristics on performance ratings.

## Additional Methodological and Conceptual Issues

There are two additional issues that must be considered with respect to developing a model that would predict performance ratings. The first is a methodological issue. Previous studies examining rating models have all used a single summary item or composite ratings as criteria (e.g.,, Hunter, 1983; Borman et al., 1991; Barrick et al., 1993; McCloy et al., 1994). Borman (1991), however, suggests that if research is concerned with specific predictor - criteria relationships it is important to use the most appropriate criteria for each predictor. While the can-do ability and knowledge measures would be expected to be predictive of ratings of task-specific skills, personality or social intelligence measures may be more predictive of ratings of interpersonal skills, and motivational constructs of ratings of effort. Using multiple criteria ratings links predictors with the most appropriate criteria, and allows an examination of the weighting policies raters use in making a single-item overall performance assessment.

The second issue is a conceptual one. Although models of performance ratings have suggested that supervisor ratings reflect both can-do and will-do factors, Murphy and Cleveland (1991) suggest that ratees are more likely to perform maximally when they are being observed by a supervisor. This may mean that when supervisors are observing, individuals do not demonstrate their typical actions and interactions, so supervisors are not able to accurately assess their typical motivation or interpersonal skills. Different groups of raters (supervisors, peers, self, subordinates) have different views of an individual's performance and, therefore, use different information in making evaluations (Borman, 1974; Murphy & Cleveland, 1991; Knapp & Campbell, 1993). Of course, the most commonly used evaluators are supervisors (Knapp & Campbell, 1993). Peer raters are a less common source of ratings, but have received considerable psychometric support (see reviews by Reilly & Chao, 1982; Kane & Lawler, 1978; Lewin & Zwany, 1976). Research comparing peer and supervisor performance ratings has suggested that there are, in fact, differences in the antecedents and structure of peer and supervisor ratings (Borman et al., 1994; Oppler, Peterson, & McCloy, 1994), and that peers may be better able to distinguish between ability and effort dimensions of performance (Klimoski & London, 1974). These findings suggest that the extent to which motivational constructs are reflected in performance ratings may well depend on the position of the rater making the evaluation.

## Hypotheses

### Path Models.

This research was designed to define and test models that assess the importance of motivational variables in predicting performance assessment. It was predicted that the motivational variables would explain a significant amount of variance in self-efficacy and effort beyond the ability and experience variables. Both distal motivational variables (achievement orientation, locus of control, and mastery beliefs), as well as a more proximal motivational variable (self-efficacy) were included in the model, and it was proposed that these variables would affect performance evaluations through observation of specific patterns of motivational effort (see Figure 1).

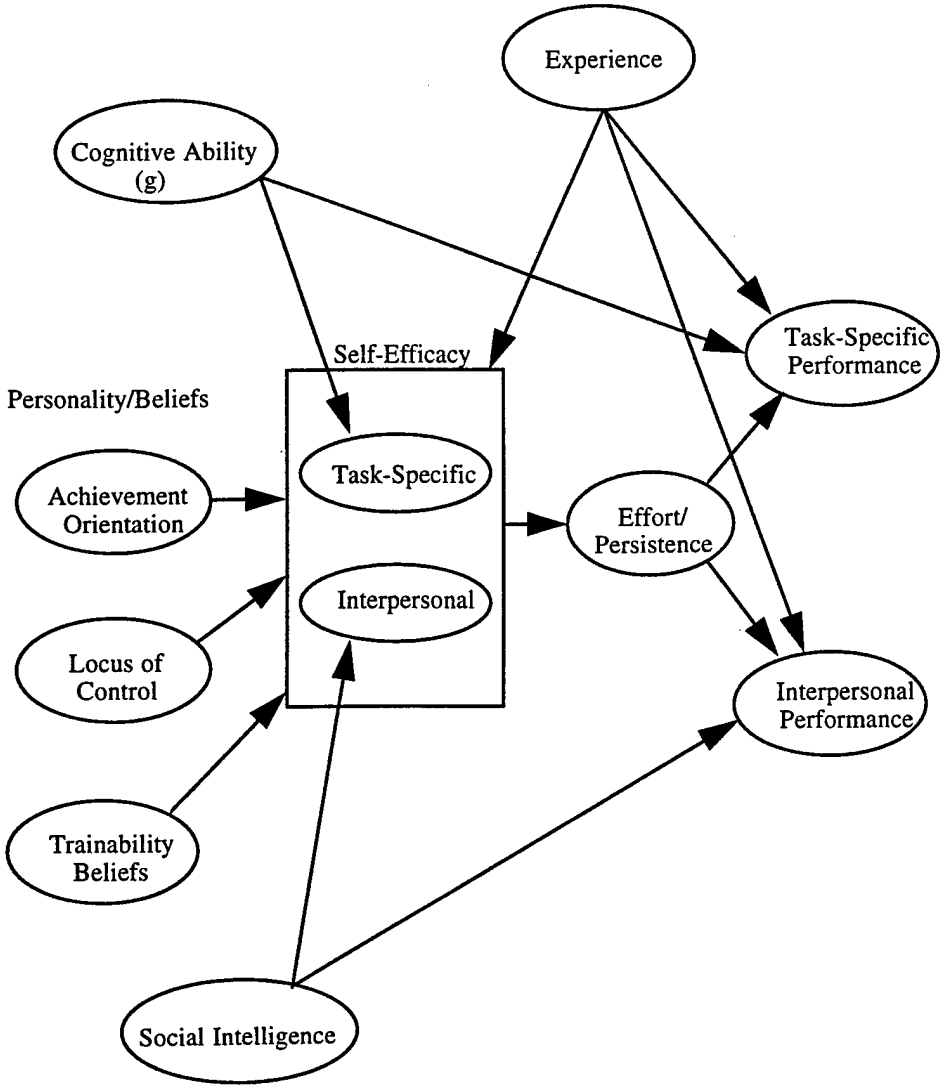


Figure 1. Proposed causal model of performance ratings.

Research has shown that peers and supervisors differ in the amount and type of information they can acquire as input to the rating process, as well as the weight they place on different dimensions of performance (Borman et al., 1994; Oppler et al., 1994; Klimoski & London, 1974). While the trainer's job is to guide and evaluate a student's demonstration of specific knowledges and skills, peers have a better opportunity to see the entire sequence of an individual's patterns of behavior and attitude over time and beyond specific task performances. This increases the amount of information they have about an individual's effort and orientation to the task, thereby strengthening the motivational and interpersonal links in the model. It was, therefore, predicted that the proposed model in Figure 1 would fit the data better using peer raters than supervisor raters. Specifically, the path coefficients for the model of peer ratings were expected to show stronger relationships between self-efficacy and effort, and between social intelligence and interpersonal performance ratings.

An Exploratory Investigation: Environmental Responsiveness.

The path model developed for this research includes motivational and ability factors in predicting the performance of students during a training program. Ford (1992) as part of his Living Systems Framework, has suggested that another critical influence on performance is the responsiveness of the environment; that is, in addition to having the requisite abilities and motivation, individuals must have a responsive environment that will allow them to progress toward the goal of performance. A responsive environment is one that provides the informational and material properties and resources that are needed for goal attainment (Ford, 1992). In the current research, the setting involved a course that was designed to provide each student with an *equally supportive* environment. While it is possible that some bias or favoritism could exist in this setting, to some degree, the effects of any bias were expected to be minimal. Given this, environmental factors were not expected to have a significant role in predicting an individual's performance.

Controlling the external environment, however, does not mean that students' *perceptions* of the supportiveness of the environment will be equivalent. Ford suggests that differences in students' perceptions of the environment, their *context beliefs,* are critical. These context beliefs may be *a consequence* of their personality, or other personal characteristics, as well as their ability to adjust their behavior to fit the environmental situation (i.e., their social intelligence). This means that these perceptions would participate in this model as a consequence of personal characteristics and antecedent to self-efficacy, rather than an external antecedent to performance.

Empirical support for this theory has not been previously provided, so a preliminary examination of this variable as a predictor to self-efficacy would be useful, and was explored separately from the path model predicting performance ratings.

`Policy Capturing Analyses.

In addition to the models examining antecedents to performance ratings, the relative importance of task-specific performance, interpersonal performance, and effort to peer and supervisor ratings of overall current performance and expected future performance was examined. The rating models reviewed earlier demonstrated that performance ratings can capture variance attributable to job knowledge and skills (Hunter, 1983; Borman et al., 1991; Borman et al., 1994; McCloy et al., 1994), as

well as interpersonal behaviors (Borman et al., 1994), and that interpersonal factors may be more important to peers than supervisors (Borman et al., 1994).

It was, therefore, predicted that, ratings of task-specific performance and interpersonal performance would each be significantly related to an overall performance rating for both peer and supervisor ratings, and that interpersonal performance ratings would have a significantly stronger regression coefficient in predicting overall performance for peers than trainers.

While effort was expected to predict ratings of task-specific performance and interpersonal performance, it was not expected to contribute significant unique variance beyond that from task-specific and interpersonal performance in predicting an overall performance rating. When raters are asked to predict an individual's overall future performance, however, as opposed to evaluating current performance, motivation would be expected to add significant unique variance beyond task and interpersonal performance; that is, demonstrated effort would be viewed as directly relevant to predicting an individual's improvement and what he or she actually *will do* in the future. It was predicted that effort would explain significant unique variance in future performance only for peers because of their greater observational opportunities (see Figure 2).

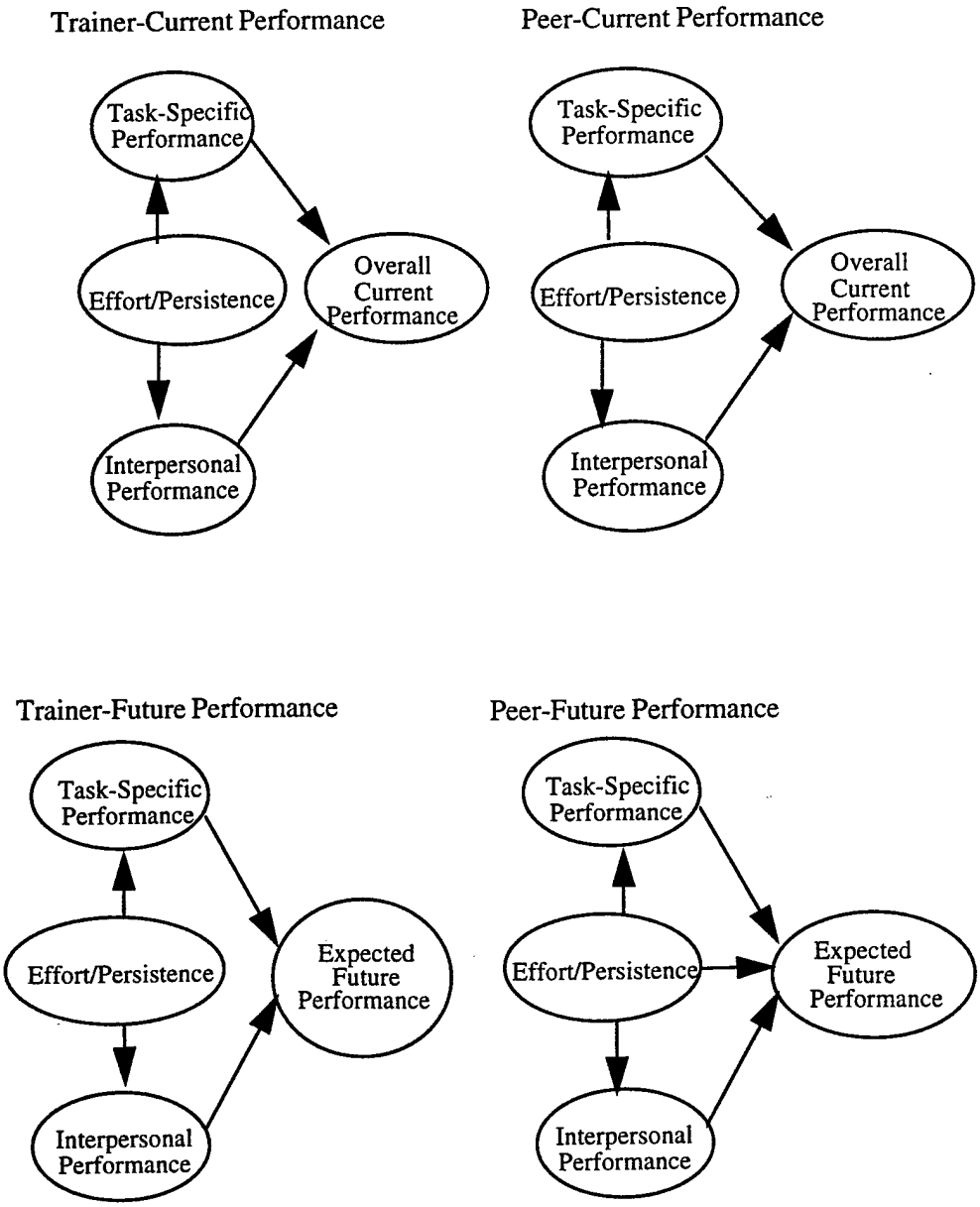


Figure 2. Four rating decision models.

## METHOD

### Subjects & Setting

Subjects were 239 enlisted male Army soldiers attending the training course for Special Forces, the Special Forces Qualification Course (SFQC) between March and October of 1995. Qualification training occurs in three phases. The first phase teaches land navigation and other tactical knowledge and skills, the second instructs each soldier in his technical specialty area (weapons, medicine, communications, or engineering), and the third teaches mission planning and preparation and requires integration of the earlier phases. Although the entire course spans several months of training, this research collected data only from the first phase of training, a segment which alone spanned 25 days.

Soldiers who successfully meet the training standards will earn the Special Forces tab; those who fail to meet the standards during any one of the three phases will either be recycled to the next class, or relieved from the course. There is, therefore, a strong selection component to the training program, so soldiers should generally be maximally motivated to perform. This may attenuate the results by restricting the range of the motivation variables and introducing skewed distributions.

Most soldiers in this sample were noncommissioned officers ("NCO's" are E-5 or higher); 58% sergeants (E-5), 16% staff sergeants (E-6). Only 26% were specialists/corporals (E-4). The majority were on active duty (70%), with the remaining 30% in the reserve. Intelligence test scores and personality and belief data were available for 151 of the 239 soldiers, so the modeling analyses that included these measures were based on this subsample of 151. No notable differences were found between the soldiers *with* intelligence, personality, and belief scores (N=151) and those *without* (N=88) on the remaining available variables (see Appendix A). The modeling analyses were, therefore, assumed to be based on a random selection.

Additional demographic data were also available for the subsample. These data indicated that soldiers ranged in age from 20 to 35, with an average age of 25, and that 87% of the soldiers described themselves as white. The soldiers had been in the Army for an average of 4.5 years, and half of them had one or more years of education beyond high school. Soldiers attending this course were selected for training based on their performance during a 21-day assessment center (Special Forces Assessment and Selection, or SFAS). At SFAS they were assessed on a variety of factors, including general intelligence, physical fitness, leadership potential, and ability to work within a team setting.

### Measures

#### Ability.

Basic cognitive ability was measured using the Wonderlic Personnel Test (Wonderlic, 1988). The range of cognitive ability scores in this sample was expected to be somewhat restricted because the General/Technical (G/T) composite of the Armed Services Vocational Aptitude Battery, which

encompasses basic verbal and math components, was used as a prescreening hurdle for placement into SFAS. Basic descriptive statistics for the Wonderlic and other predictors in the model can be seen in Table 1.

**Table 1. Descriptive Statistics for Predictors**

| | Mean (SD) | Range | Skew[1] | Coeff Alpha |
|---|---|---|---|---|
| **Intelligence** | | | | |
| Wonderlic | 25.41 (5.49) | 10-45 | 2.57** | NA[2] |
| Overall Social Intelligence | 3.77(0.37) | 2.8-4.9 | 2.47** | .87 |
| Social Perceptiveness | 3.75(0.42) | 2.7-4.9 | 2.35** | .87 |
| Behavioral Flexibility | 3.81(0.47) | 2.4-5.0 | -0.34 | .78 |
| **Motivation** | | | | |
| Achievement Orientation | 13.51(4.97) | 1-23 | -2.14* | NA[2] |
| External Locus of Control | 6.95(3.73) | 0-18 | 3.21** | NA[2] |
| Mastery Beliefs | 3.96(0.76) | 1.2-5.2 | -4.47** | .73[3] |
| Self Efficacy - Interpersonal | 4.01(0.79) | 1.9-5.0 | -3.14** | .91 |
| Self Efficacy - Tactical | 4.04(0.60) | 2.1-5.0 | -3.00** | .90 |

[1] Note: *$p<.05$, **$p<.01$

[2] Not Available: Reliability information not available.

[3] Reliability calculated as internal consistency of subscales

Basic social intelligence (SOCIAL) was measured by a 32-item background measure developed by Gilbert, Connelly, Mumford, and Zaccaro (1992). This measure included two subscales of social intelligence (social perceptiveness (SOCPER) and behavioral flexibility (BEHFLEX), which were combined for this research into a single scale. Example items from scales are as follows: (a) social perceptiveness - "How long does it take you to figure out when someone is upset?"; (b) behavioral flexibility - "How difficult has it been to be polite to people you dislike when meeting in a social situation?" The scale showed high internal consistency (alpha=.87), although scores evidenced a restricted range (minimum value = 2.8 on a scale from 1-5, where 1 represents a low score), as well as significant positive skew ($z=2.47$, $p<.01$).

Experience.

A composite experience score (EXPER) was created using information obtained from records. The goal in creating this variable was to assign students a value on a continuous scale that represented a composite of their relevant experience. Three variables were identified by previous research as particularly relevant to this situation: the student's rank, prior branch type, and ranger qualification status (Brooks, 1997). These measures were expected to be related to both task-specific aspects of an individual's training as well as interpersonal aspects.

Due to a considerable overlap in the program of instruction between Ranger training and training for Special Forces, Ranger qualification status most strongly predicts SFQC success. Lappin (1995) found that Ranger-qualified students were four times more likely to graduate from the SFQC than students who did not have combat experience in their prior branch and are not yet NCOs. Of the current subsample, 47 (31%) were Ranger- qualified. These subjects were assigned a "3", the highest value, for the "Experience" variable.

The other factors that were shown to be related to higher success rates were higher rank (E-5 or higher) and prior combat arms branch types. Based on success rates found by Lappin (1995), soldiers who had both of these experience factors were assigned a "2", and soldiers who had only one of the two factors were assigned a "1". Soldiers who did not have any of these experience factors were assigned a "0". The distribution of this variable was as follows: 31% high experience (3), 31% moderate experience (2), 25% low experience (1), and 13% no experience.

Achievement Orientation.

Achievement orientation (ACHO) was measured using the 38-item Achievement Scale from the Adjective Checklist (ACL) (Gough & Heilbrun, 1983). The ACL has been very widely used and has demonstrated acceptable alpha coefficients (.76 for males) and test-retest reliabilities (.65 for males) (Gough & Heilbrun, 1983). Using this to measure achievement orientation was beneficial in that it required only 15-20 minutes for completion, whereas other measures that included achievement orientation would have required an hour or longer (e.g.,, Personality Research Form, Jackson, 1984). Subjects were instructed to select all words that described them, and achievement scores were calculated by subtracting the number of negative achievement-oriented words selected from the number of positive ones selected. Subjects showed a broad range of scores (1-23), but the scores showed some negative skew. Higher scores on this measure indicated a higher orientation toward achievement.

Locus of Control.

Locus of control (EXTLOC) was measured using Rotter's (1966) 29-item forced-choice Internal-External questionnaire (6 filler items). It has been used extensively and has demonstrated reliabilities ranging from .6 to .9 (Jonassen & Grabowski, 1993). While other scales exist that include additional attributional dimensions, the focus of Rotter's measure was most appropriate for the hypotheses of this research; Rotter's scale measures a generalized expectancy to perceive outcomes as either contingent upon or external to one's own actions. Further, Rotter's scale is appropriate for adult populations, while several of the others were developed specifically for child populations. Scores in this sample ranged from 0 to 18, but evidenced significant positive skew. Higher scores indicated an external locus of control.

Mastery Beliefs.

A 29-item agree/disagree measure was developed to measure helplessness/mastery beliefs for this research, using items such as those described by Chiu, Hong, and Dweck (1994) (see Appendix B). This measure was developed because no appropriate established measure of this construct was found. Items were designed to measure the extent to which an individual believes the abilities and skills needed for performance in the setting are innate or are trainable, and covered the following domains: tactical/field,

interpersonal, leadership, physical, and technical/academic. Mastery responses to the items, responses agreeing that the attributes in a domain were trainable, were counted within each domain, then averaged across domains for an overall mastery value (MASTERY). Of a possible range from 0 to 5.2, scores in this sample ranged from 1.2 to 5.2, with higher scores reflecting higher levels of mastery orientation, and were significantly negatively skewed. The summary measure had a moderately high internal consistency coefficient across the five subscales (alpha=.73), and intercorrelations among the subscales ranged from a low of .16 between the social and physical domains, to a high of .52 between tactical and academic domains. The measure showed convergent validity in its significant correlation with locus of control ($r = -.23$, $p<.01$), which suggested that higher mastery beliefs were related to an internal locus of control. The measure was not significantly related to achievement orientation.

Self-efficacy.

Task and interpersonal self-efficacy measures (TASKEFF, SOCEFF) were developed for this study through consultation with subject matter experts from the training program (see Appendix C). The measures were modeled after those used in Zaccaro, Zazanis, Diana and Greathouse (1994), presenting a series of items that ask subjects to indicate, on a 5-point scale, their level of confidence in their ability to complete situation-relevant tasks. In order to minimize the effects of social desirability, the difficulty of the tasks used in the items was varied from extremely easy to nearly impossible. Students were instructed that we were interested in examining how well they were able to accurately assess their training progress. Responses were averaged to create task and interpersonal self-efficacy scores. Ranges for both variables were slightly truncated at the low end (minimum scores were around 2.0), averages were high (4 .0 for both), reflecting generally high levels of self-efficacy, and distributions were significantly negatively skewed. Coefficient alpha reliabilities were high for both scales: .90 for task self-efficacy, and .91 for interpersonal self-efficacy.

Ratings.

Seven ratings were obtained from peers and trainers: interpersonal performance, effort/persistence, performance in three task-specific domains (physical fitness, leadership, military tactics), overall current performance, and predicted future performance (see Appendix D). These scales were developed for the training company using the theoretical tenets of this research project. The rating scales for the five specific ratings provided the rater with behavioral indicators of performance, or factors to consider when making their decisions. Both trainers and peers used identical rating forms.

Ratings for physical fitness, leadership, and tactical skills were averaged to form the overall task performance rating. On a 5-point rating scale, averages for the peer ratings were somewhat higher than the scale midpoint, ranging from 3.34 (task performance) to 3.70 (future performance), and, all ratings except those for task performance showed significant negative skew (see Table 2). The distributions of trainer ratings were normal, and they showed averages only slightly above 3.0. The means for the peer ratings were significantly greater than for trainer ratings (all t's, $p<.01$), and the variances for the trainer ratings were significantly greater than those for the peers (all F's, $p<.01$). Overall, some negative skew in the ratings had been expected, given that the sample was selected for training based on their potential for success in this setting.

For each student, 8-12 peer ratings and a single trainer rating were obtained. Interrater reliabilities for the peer ratings were calculated using two methods. The first used the coefficient alpha to estimate internal consistency among the peer raters within each group. These reliabilities were high, with coefficients ranging from .75 for interpersonal performance to .95 for task performance. The second method averaged ratings for the odd and even raters, then correlated these averages for each ratee. These split-half reliability estimates were also high, ranging from .71 (interpersonal) to .85 (task and overall performance).

Interrater reliabilities could not be calculated for trainer ratings because only one set of trainer ratings was available for each student; however, a meta-analysis of supervisor rating studies found that the interrater correlation for "perfect" measures is virtually constant at .60 across all studies (King, Hunter, & Schmidt, 1980). Research comparing the reliabilities of peer and supervisor ratings has suggested that the reliability of a single supervisor was equivalent to the reliability of the average of 2.25 peer raters; that is, 2.25 peer raters were required to achieve the reliability of a single supervisor (Oppler, Pulakos, & Borman, 1992). Peer ratings in the current analyses, then, would be expected to have higher reliability than the supervisor ratings.

**Table 2. Rating Descriptives and Reliabilities**

| | Mean (SD) | Range | Skew[1] | Coeff Alpha | Split Half |
|---|---|---|---|---|---|
| **Peer Ratings** | | | | | |
| Effort | 3.68(0.66) | 1.3-5.0 | -3.35** | .81 | .73 |
| Interpersonal | 3.46(0.62) | 1.2-4.8 | -3.33** | .75 | .71 |
| Task | 3.34(0.73) | 1.1-4.8 | -1.23 | .95 | .85 |
| Overall | 3.47(0.76) | 1.2-5.0 | -2.36** | .88 | .85 |
| Future | 3.70(0.71) | 1.2-5.0 | -3.43** | .83 | .77 |
| **Trainer Ratings** | | | | | |
| Effort | 3.21(1.06) | 1-5 | -0.69 | NA[2] | NA |
| Interpersonal | 3.09(1.07) | 1-5 | -0.86 | NA | NA |
| Task | 3.00(1.02) | 1-5 | -0.17 | NA | NA |
| Overall | 3.01(1.07) | 1-5 | -1.43 | NA | NA |
| Future | 3.23(1.08) | 1-5 | -1.64 | NA | NA |

[1]Note: *$p$<0.5 and **$p$<.01

[2]Not Available: Reliability information not available because only one trainer was available to produce ratings.

Procedure

A scannable survey measuring social intelligence, achievement orientation, locus of control, and mastery beliefs was administered by a researcher during in-processing for the course, before any training was conducted. Soldiers were instructed that the survey was voluntary, and that their responses would be used for research purposes only; it was not part of their training at SFQC. Fewer than 2% of the soldiers elected to turn in surveys without completing them.[1] Task and social self-efficacy were collected by a researcher in the field around day 15 of the course. This timing was chosen because it was after the classroom training and practice exercises (so all soldiers were familiar with the terms and activities described in the measure), but before the series of graded performance exercises. Again, soldiers were instructed that the survey was voluntary, and that their responses would be used for research purposes only; fewer than 1% of the soldiers elected not to fill out the survey. Finally, peer and trainer ratings were collected by training program personnel in the field after the final performance exercises were completed. Wonderlic and experience data were obtained from course records.

[1] High response rates were not surprising because soldiers' options during survey administrations were either to fill out the survey or to sit and wait for the next activity while others filled out the survey.

## RESULTS

Zero Order Correlations

Predictor Intercorrelations.

Correlations among the ability and motivational measures were computed and examined for evidence of convergent and discriminant validity (see Table 3). Appropriately, the Wonderlic measure was unrelated to social intelligence and the personality measures. An unexpected negative correlation between Wonderlic scores and efficacy, however, suggested that higher cognitive ability was related to lower task and interpersonal efficacy. The newly created mastery beliefs measure showed convergent validity in its significant correlation with an internal locus of control, and divergent validity in its lack of relationship with Wonderlic and social intelligence measures. Surprisingly, achievement orientation showed no relationship with either mastery beliefs or locus of control.

**Table 3. Predictor Intercorrelations**

| | WOND | SOCIAL | EXPER | ACHO | EXTLOC | MASTERY | TASKEFF |
|---|---|---|---|---|---|---|---|
| SOCIAL | .02 | | | | | | |
| EXPER | -.14 | -.10 | | | | | |
| ACHO | -.01 | .09 | -.09 | | | | |
| EXTLOC | .07 | -.26** | -.01 | -.03 | | | |
| MASTERY | -.16 | .09 | .04 | -.07 | -.23** | | |
| TASKEFF | -.29** | .13 | .34** | -.03 | -.07 | .12 | |
| SOCEFF | -.27** | .25** | .20* | .05 | -.16* | .03 | .62** |

Note: *$p<.05$, and ** $p<.01$

As would be expected, the efficacy measures were correlated (r=.62, $p<.01$), but were not redundant; respondents did discriminate between task and interpersonal self-efficacy. Appropriately, interpersonal efficacy was significantly related to social intelligence, while task efficacy was not. Interpersonal efficacy was also significantly related to an internal locus of control, while task efficacy did not show this relationship.

Predictor-Criteria Correlations.

First-order correlations between ratings and the ability and personality motivational measures were small in magnitude and nonsignificant (see Table 4). Task self-efficacy and experience, however, evidenced strong significant relationships with both peer and trainer ratings, although interpersonal self-efficacy was unrelated to the rating measures. Peer ratings showed more discriminant validity for task self-efficacy than did trainer ratings, with a correlation of .31 ($p<.01$) between task self-efficacy and peer task rating, and a nonsignificant correlation between task self-efficacy and peer social rating; corresponding correlations for trainers were .31 and .23 ($p<.01$) , respectively.

Wonderlic -Efficacy Relationship.

Wonderlic scores showed unusual zero-order correlations with self-efficacy and performance ratings; both relationships were expected to be significant and positive but the zero-order correlation with self-efficacy was significant and negative, and that with performance was nonsignificant. The significant negative correlation between Wonderlic and self-efficacy indicated that students who were high in cognitive ability expressed lower confidence that they had the skills and abilities required to accomplish the task-specific and interpersonal requirements of performance; scores on cognitive ability were not, however, related to performance ratings. The restriction in range for cognitive ability in this sample may have attenuated some correlations.

**Table 4. Predictor-Criteria Correlations**

| | Predictors | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | WOND | SOCIAL | EXPER | ACHO | EXTLOC | MASTERY | TASKEFI | SOCEFF |
| PEER | | | | | | | | |
| Effort | .01 | -.06 | .36** | .07 | .07 | -.05 | .14 | .00 |
| Social | .04 | -.09 | .30** | .00 | .06 | -.02 | .07 | -.02 |
| Task | -.04 | .00 | .61** | -.05 | .08 | -.03 | .31** | .11 |
| Overall | .02 | -.05 | .49** | -.03 | .09 | -.04 | .23** | .05 |
| Future | .01 | -.04 | .45** | .00 | .11 | -.05 | .21** | .07 |
| TRAINER | | | | | | | | |
| Effort | -.08 | -.01 | .29** | -.03 | .07 | -.08 | .19* | .06 |
| Social | -.08 | .07 | .36** | -.03 | .10 | -.04 | .23** | .10 |
| Task | -.05 | .04 | .51** | -.06 | .05 | -.01 | .31** | .13 |
| Overall | -.07 | .01 | .43** | -.05 | .05 | -.08 | .19* | .08 |
| Future | .02 | .00 | .39** | -.05 | .05 | -.10 | .09 | -.03 |

Note: *$p<.05$, and ** $p<.01$

Criteria Intercorrelations.

The effort and performance ratings showed a rater effect, with higher correlations within the rater groups than between peer and trainer ratings for a given dimension (see Tables 5 and 6). Intercorrelations for effort, task, and interpersonal ratings within a rating group were extremely high (an average of .79 for peers and .78 for trainers). They were not combined, however, given their differentiated role in the hypotheses. These high intercorrelations may attenuate results that differentiate among the ratings. For peers, the largest differentiation between performance dimensions, was between task and interpersonal performance; for trainers, the task, interpersonal, and effort intercorrelations were all in the .76-.80 range.

**Table 5. Intercorrelations for Peer and Trainer Ratings**

*Peer Evaluations*

| | Effort | Interp | Task | Overall |
|---|---|---|---|---|
| **Interp** | .81 | | | |
| **Task** | .82 | .74 | | |
| **Overall** | .89 | .83 | .95 | |
| **Future** | .90 | .86 | .91 | .96 |

*Trainer Evaluations*

| | Effort | Interp | Task | Overall |
|---|---|---|---|---|
| **Interp** | .77 | | | |
| **Task** | .80 | .77 | | |
| **Overall** | .80 | .76 | .90 | |
| **Future** | .78 | .76 | .86 | .89 |

Peer and trainer raters showed the strongest agreement when rating task performance and overall current performance, with correlations of .82 and .80, respectively; they showed the least agreement for ratings of effort and interpersonal performance, although the correlations were still high, with correlations of .63 and .65, respectively (see Table 6). This is consonant with the hypothesis that peers and trainers have different opportunities to view behaviors related to effort and interpersonal skills.

**Table 6. Correlations Between Peer and Trainer Ratings**

*Trainer Ratings*

| *Peer Ratings* | Effort | Interp | Task | Overall | Future |
|---|---|---|---|---|---|
| **Effort** | .63 | .64 | .67 | .66 | .67 |
| **Interp** | .58 | .65 | .59 | .61 | .66 |
| **Task** | .67 | .71 | .82 | .81 | .77 |
| **Overall** | .69 | .74 | .80 | .80 | .79 |
| **Future** | .68 | .74 | .77 | .76 | .77 |

In the sections that follow, LISREL8 with a maximum likelihood estimation method was used to examine model links and to compare the overall fit of the peer and trainer models. Following that, multiple regression was used to examine the relevance of the environmental responsiveness variable, and to test the policy-capturing hypotheses. All path coefficients ("b" or "beta") reported here refer to standardized coefficients.

Path Models Predicting Performance Ratings

Initial results from LISREL8 for both the peer and trainer models indicated a number of nonsignificant path coefficients (see Figures 3 and 4). Model fit statistics suggested a poor fit of the original model to the data, and modification indices recommended links that could be added to improve model fit (see Tables 7 and 8).

Modifications to the models.

The initial peer model showed a chi-square of 122.10 (23df), which was significant, with an adjusted goodness of fit index (AGFI) of .66. The initial trainer model showed a chi-square of 123.60 (23df), which was significant, with an AGFI of .65. Modification indices for both models suggested first adding a link from task self-efficacy to interpersonal self-efficacy. While these two constructs may be somewhat reciprocal, for the purposes of this model it is probable that a student's confidence in accomplishing the task at hand would tend to increase his confidence in the social interactions he engages in during task performance. When the path from task to interpersonal self-efficacy was added to the peer model, chi-square (22df) improved to 55.76, $p<.01$, with an AGFI=.81; for trainers, chi-square (22df) improved to 57.68, $p<.01$, with an AGFI=.80.

**Table 7. LISREL Modifications to the Peer Model**

| Links Added to the Peer Model | $\chi^2$(df) | Significance | AGFI |
|---|---|---|---|
| 1. Model as originally specified | 122.10 (23) | $p<.01$ | .66 |
| 2. Task Self-efficacy to Interpersonal Self-efficacy | 55.76 (22) | $p<.01$ | .81 |
| 3. Experience to Effort | 39.46 (21) | $p<.01$ | .85 |
| 4. Task Self-efficacy to Task Performance | 30.23 (20) | $p<.05$ | .88 |
| 5. Correlated residuals between Task and Interpersonal Performance | 22.31 (19) | $p>.05$ | .91 |
| 6. Final Modified Model - nonsig links out | 18.78 (15) | $p>.05$ | .92 |

**Table 8. LISREL Modification to the Trainer Model**

| Links Added to the Trainer Model | $\chi^2$(df) | Significance | AGFI |
|---|---|---|---|
| 1. Model as originally specified | 123.60 (23) | $p<.01$ | .65 |
| 2. Task Self-efficacy to Interpersonal Self-efficacy | 57.68 (22) | $p<.01$ | .80 |
| 3. Correlated residuals between Task and Interpersonal Performance | 31.32 (21) | $p>.05$ | .88 |
| 4. Experience to Effort | 21.77 (20) | $p>.05$ | .91 |
| 5. Social Intelligence to Task Performance | 16.53 (19) | $p>.05$ | .93 |
| 6. Final Modified Model - nonsig links out | 10.96 (12) | $p>.05$ | .94 |

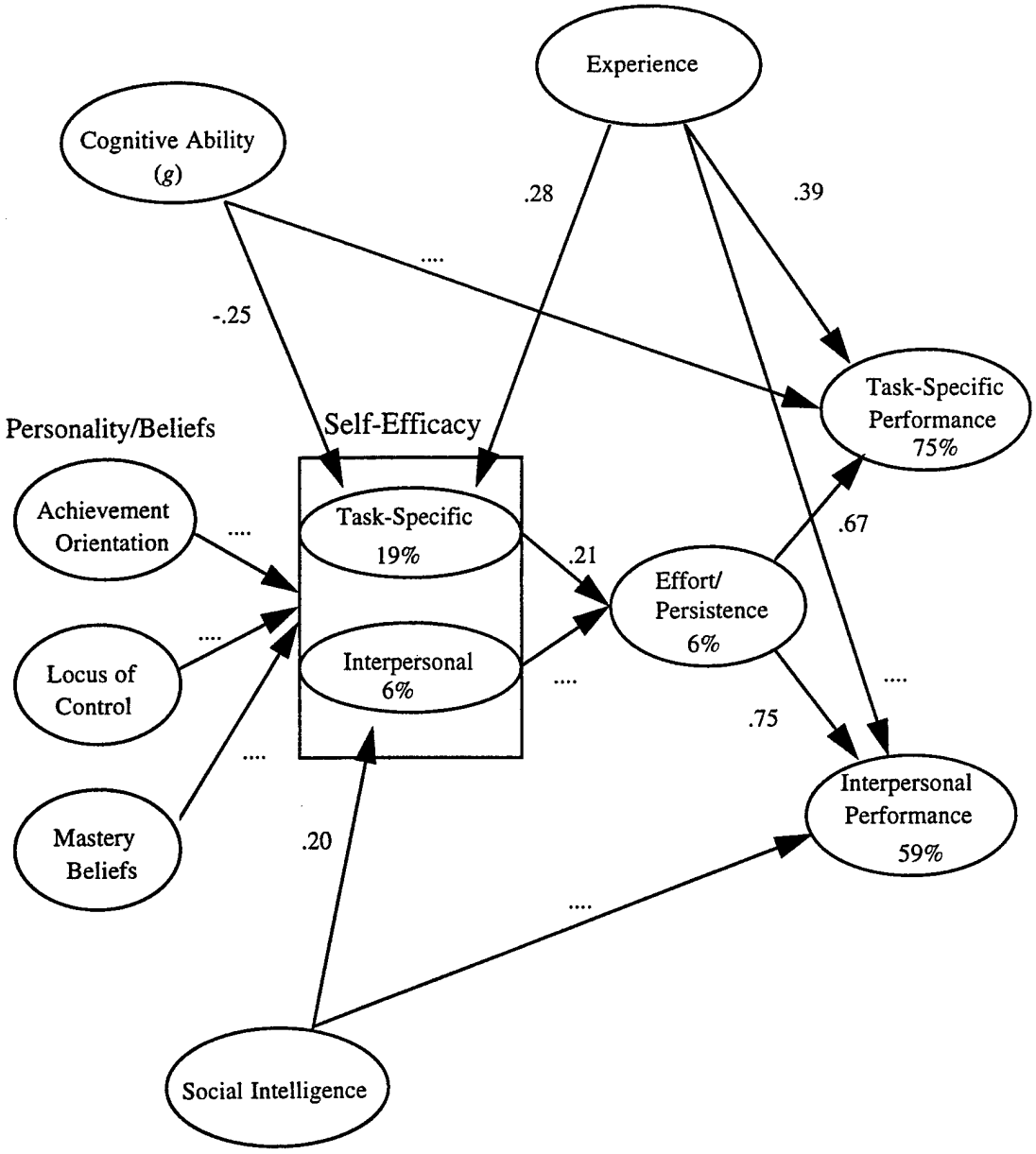


Figure 3. Initial results for the model predicting peer ratings of performance.

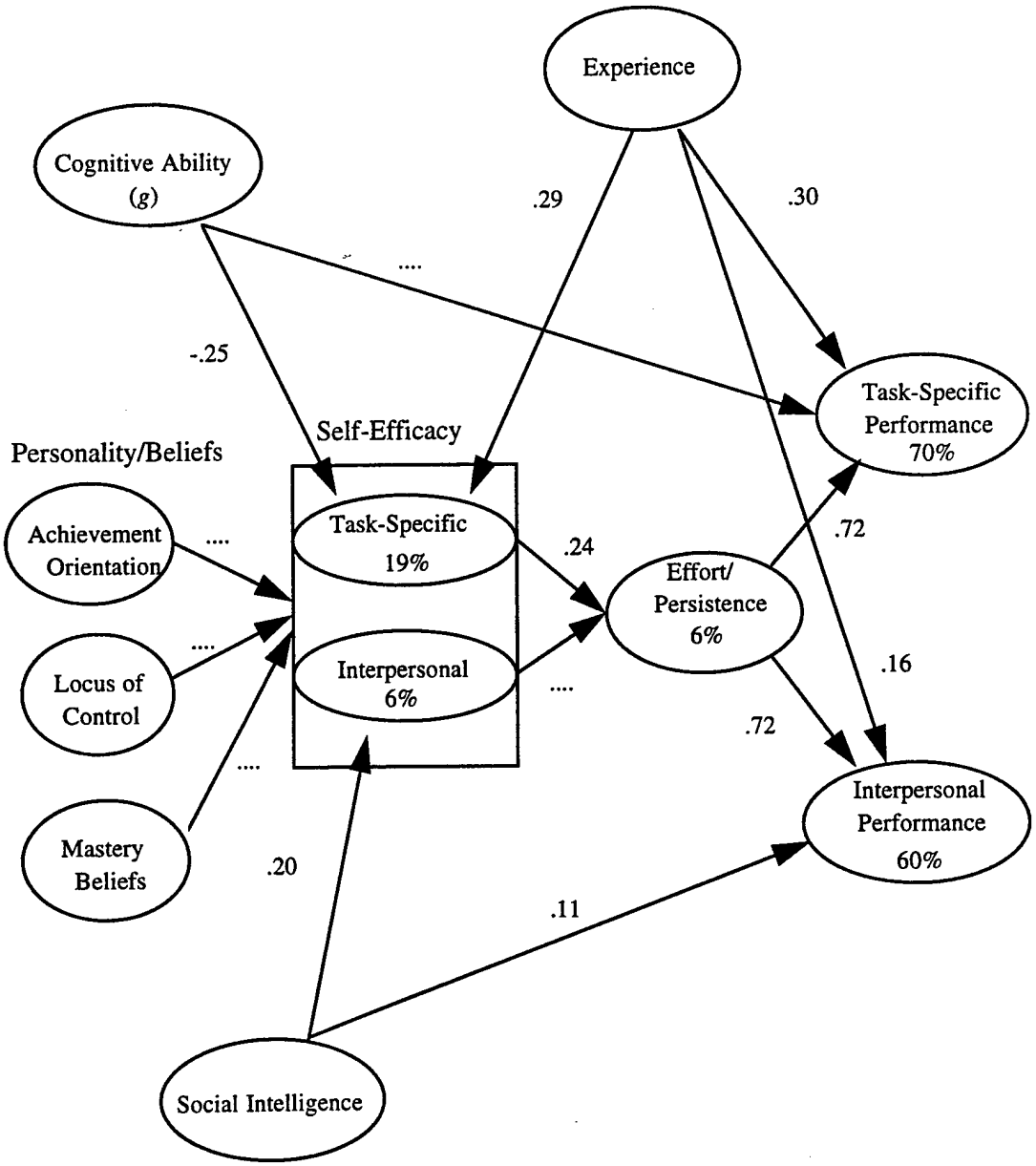


Figure 4. Initial results for the model predicting trainer ratings of performance.

For the peer model, the second modification recommended was a link from experience to effort, suggesting that students with higher levels of experience were directly perceived by peers as exerting more effort. Adding this link improved chi-square (21df) to 39.46, $p<.01$, with an AGFI of .85. The third modification recommended was a direct link from task self-efficacy to task performance, improving chi-square (20df) to 30.23, $p<.05$, with an AGFI of .88. The final modification recommended by LISREL8 was for correlation between the task and interpersonal performance variables. Given that these measures have the same format and were collected at the same time, this modification is reasonable. The final change improved chi-square (19df) to 22.31, which was not significant, and the model had an AGFI of .91. The final modified model demonstrated a good fit to the data, and can be seen in Figure 5.

For the trainer model, the second modification recommended was to allow for correlations between task and interpersonal performance. This improved chi-square (21df) to 31.32, which was not significant, with an AGFI of .88. The third modification recommended was the link between experience and effort, improving chi-square (20df) to 21.77, not significant, with an AGFI of .91. The final link recommended was from social intelligence to task performance, a link which is not necessarily logical given that the zero order correlation between the two variables was not significant; this link was not added. The final modified model demonstrated a good fit to the data, and can be seen in Figure 6.

Task-specific performance.

Ratings of task-specific performance were expected to be predicted by effort, general cognitive ability, and experience. This was partially supported. For both peers and trainers, the paths from effort and experience to task specific performance were significant; however, the path from cognitive ability was not (for peer effort, b=.68; for peer experience, b=.34; for trainer effort, b=.72; for trainer experience, b=.30; for all b, $p<.01$).

For peers, 80% of the variance in task specific performance was explained by the direct effects, and for trainers, 73% was explained. The additional link from task self-efficacy to performance is particularly interesting because it suggests that there is information about self-reported competence and its attendant motivational consequences that is captured by peer ratings but is not captured by the trainer ratings.

Interpersonal performance.

Interpersonal performance was expected to be predicted by ratings of effort, social intelligence, and experience. This was also partially supported. For the trainer model, both effort and experience had significant direct effects (b=.72, $p<.01$; b=.16, $p<.05$), but social intelligence did not; overall, 61% of the variance in interpersonal performance was explained. For the peer model, only effort had a significant direct effect (b=.77, $p<.01$), with a total of 59% of the variance explained.

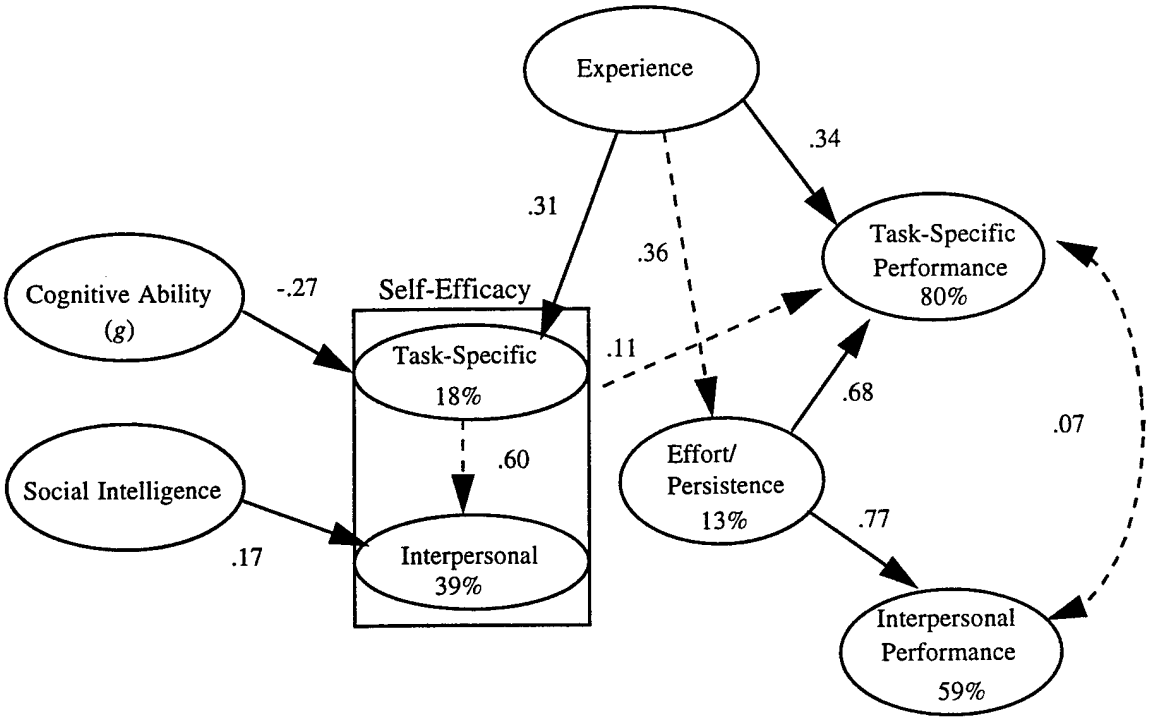


Figure 5. Final model predicting peer ratings of performance.

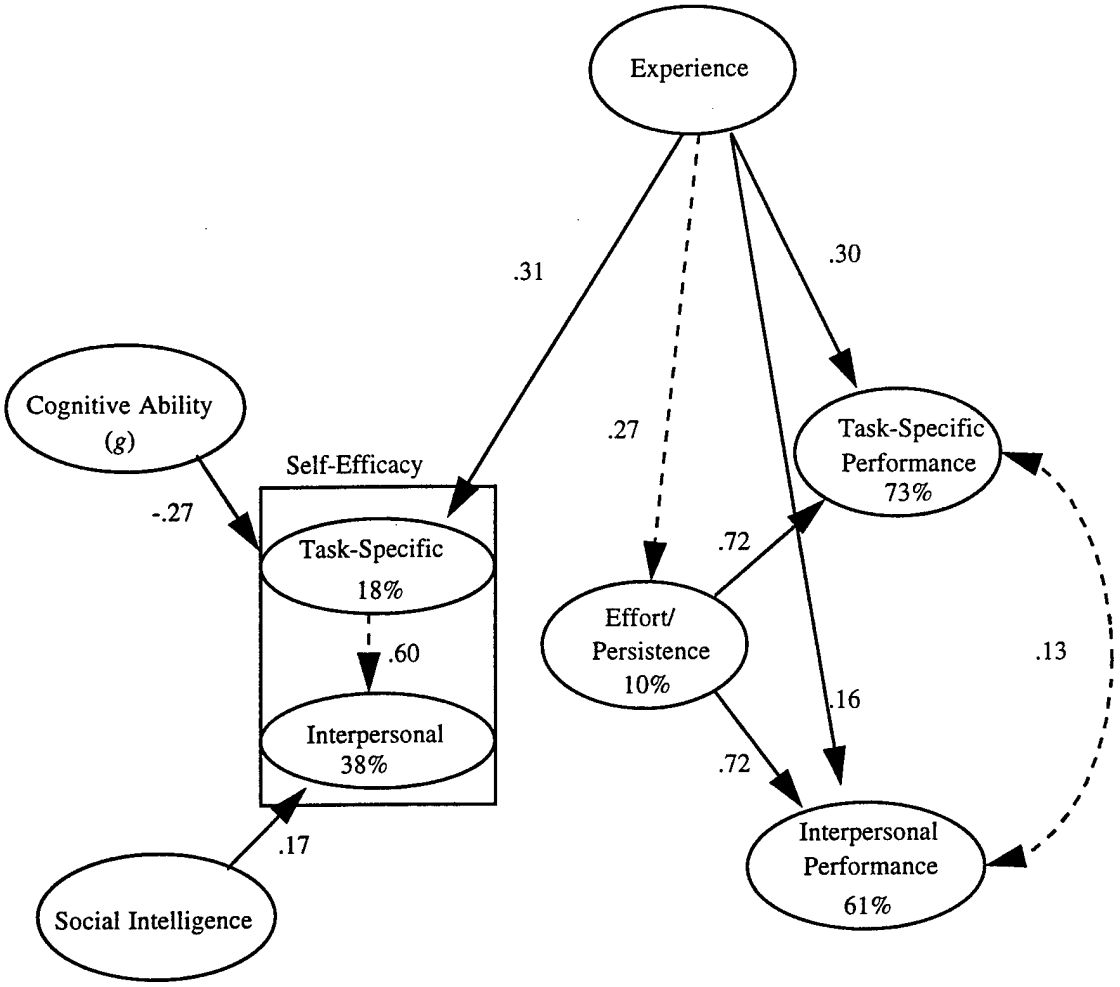


Figure 6. Final model predicting trainer ratings of performance.

Predictors of Effort.

Ratings of effort were expected to be predicted by task and interpersonal self-efficacy. This was not supported; while the path from task-specific efficacy to effort was significant for both models initially (b=.21, $p<.05$ for peers, b=.24, $p<.05$ for trainers), once experience was added as a direct predictor of effort, the link between task self-efficacy and effort became non-significant. Further, the path from interpersonal efficacy was nonsignificant for both models. Together, the two efficacy variables explained 6% of the variance in peer-rated and trainer-rated effort. When experience was added as a predictor, results showed that experience had a significant direct effect on effort both for peers (b=.36, $p<.01$) and trainers (b=.27, $p<.01$). The direct effect of experience on effort indicated that prior relevant experience *directly* influenced perceptions of the amount of effort expended by the student.

Predictors of Self-efficacy.

It was expected that the self-efficacy constructs would be significantly predicted by the personality/belief measures, with task-specific efficacy also predicted by cognitive ability and experience, and interpersonal efficacy also predicted by social intelligence. This was partially supported. None of the personality/belief measures showed significant paths to either efficacy measure. Social intelligence, however, had a significant direct relationship with interpersonal efficacy (b=.17, $p<.01$), and cognitive ability and experience had significant direct relationships with task-specific efficacy (b=-.27, $p<.01$ for cognitive ability; b=.31, $p<.01$ for experience). As evidenced in the zero-order correlations, the relationship between cognitive ability and task-specific efficacy was unexpectedly negative, suggesting that higher cognitive ability scores were predictive of lower self-efficacy scores. Prior to adding the link from task self-efficacy to interpersonal self-efficacy, 19% of the varince in task-specific efficacy and 6% of the variance in interpersonal efficacy was explained by model variables. When task self-efficacy was included as a predictor for interpersonal efficacy and the nonsignificant links were removed, this changed to 18% of the variance explained in task self-efficacy and 38% of the variance explained in interpersonl.

Comparing peer and trainer models.

The proposed model was predicted to fit the data better using peer raters than using supervisor raters, and path coefficients between efficacy and effort and social intelligence and interpersonal performance ratings were predicted to be significantly stronger for peers than supervisors. These hypotheses were not supported. With respect to the path coefficients, reviewing the coefficients between the two models shows that they were nearly identical in magnitude (review Figures 5 and 6); the difference between regression coefficients for the efficacy-effort path in the initial model was .03, a difference which was nonsignificant at the .05 level; the social intelligence-interpersonal performance path was nonsignificant for both models.

While the magnitudes of the regression coefficients between self-efficacy and effort were equivalent for peers and trainers, calculation of the 95% confidence intervals using multiple regression suggested that the coefficients for the trainer model were more questionable than those for the peer model. For peers, the confidence interval for effort regressed on task self-efficacy ranged from just above zero (.001) to .42. For trainers, this confidence interval was much wider and fell

clearly below zero, ranging from -.09 to .59, due to a greater amount of total error variance in prediction of effort for trainers. A similar pattern was seen for the path between effort and social self-efficacy for the two models; the interval for peers ranged from -.03 to .29, and for trainers it ranged from -.16 to .35.

With respect to the goodness of fit of the models, the hypothesis that the proposed model would fit the data better using peer raters than supervisor raters was not supported. Both models appeared to evidence about the same goodness of fit, both prior to modifications, where chi-square (23 df)=122.10 for the peers, AGFI=.66, and chi-square (23 df)=123.60 for the trainers, AGFI=.65, as well as after the modifications, where chi-square (15 df)=18.78 for the peers, AGFI=.92 and chi-square (12 df)=10.96 for the trainers, AGFI=.94.

A Potential Antecedent to Self-Efficacy: Environmental Responsiveness

The path models in this research included motivational and ability factors in predicting the performance of students during a training program. As suggested by Ford (1992), the possibility was examined that differences in students' perceptions of the environment, or *context beliefs,* existed that were a consequence of their personality and ability to adjust their behavior to fit the environmental situation (i.e., their social intelligence) and were antecedent to self-efficacy.

A 10-item scale measuring context beliefs about environmental responsiveness was created for this purpose based on the construct defined by Ford (1992). Items addressed the student's perception of his trainers, course resources, his fellow team members, and personal factors (see Appendix E), and the scale evidenced an acceptable reliability (coeff alpha=.69).

Perceptions of course environment were included in a path model predicting task-specific and interpersonal self-efficacies (see Figure 7). Regression path analysis results showed a strong direct effect of environmental responsiveness on both interpersonal self-efficacy (b=.24, $p<.01$) and task self-efficacy (b=.30, $p<.01$). Although the personality measures did not predict environmental responsiveness, social intelligence did have a significant direct effect (b=.24, $p<.01$). Social intelligence and cognitive ability also maintained direct paths to interpersonal efficacy. With the addition of context beliefs, the amount of variance explained in interpersonal self-efficacy increased from 13% to 16%, and the amount of variance explained in task self-efficacy increased from 16% to 22%.

The strength of these results clearly suggests the utility of further investigation of this construct. It may, in fact, be necessary to divide the construct into multiple subdimensions; correlations in Table 9 show divergent relationships for two subdimensions of environment, "team environment" and "course environment," when correlated with locus of control, social intelligence, and task and interpersonal self-efficacy.

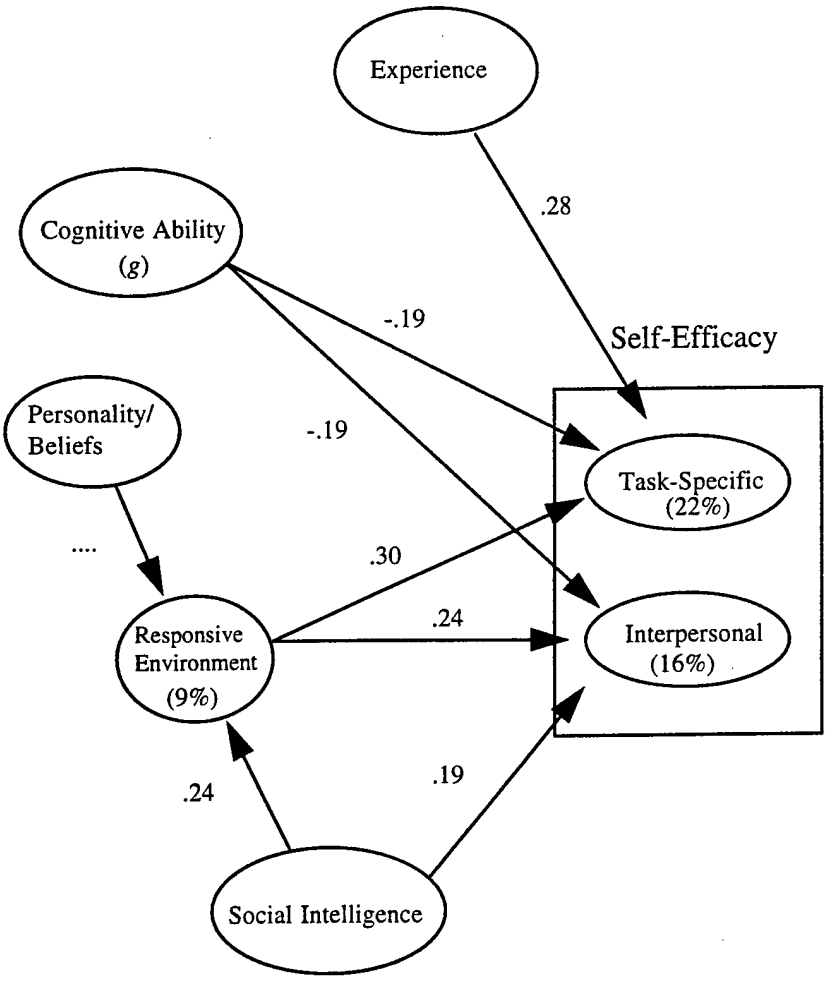


Figure 7. Causal model using responsive environment to predict self-efficacy

**Table 9. Correlations of Course Environment Scales with Other Predictors**

| | TOTAL ENV | TEAM | COURSE |
|---|---|---|---|
| WOND | -.18* | -.10 | -.17* |
| EXPER | -.05 | -.05 | .03 |
| ACHO | .07 | .02 | .07 |
| EXTLOC | -.16 | -.23** | -.06 |
| MASTERY | .12 | .15 | .00 |
| SOCIAL | .28** | .31** | .15 |
| TASKEFF | .25** | .12 | .23** |
| SOCEFF | .29** | .17* | .25** |

Note: *$p$<.05, and ** $p$<.01

These results suggested that individuals' perceptions of the responsiveness of their peers ("TEAM") may be a consequence of their ability to perceive and interact with their social environment as well as having an internal locus of control. In addition, these perceptions may be a predictor of their confidence that they have the skills required to successfully perform interpersonally. Individuals' perceptions of the responsiveness of the course structure and trainers ("COURSE") were related to their confidence in their abilities to perform both in interpersonal and task domains.

Summary

The results of the path models demonstrated that motivation, as measured by effort, is a critical antecedent to performance ratings. Neither the personality measures nor interpersonal self-efficacy, however, played a significant role in the models, and task self-efficacy was significant only in the peer model and only in predicting task-specific performance directly. In addition, results did not find the predicted differences between peers and trainers in the determinants of task-specific and interpersonal performance ratings. The post-hoc investigation of context beliefs offered preliminary information regarding potential antecedents to self-efficacy in these models.

The following analyses will determine whether there are differences in the importance given to motivation, task, and interpersonal factors by peers and trainers when rating a student's *overall* or global performance.

Policy-capturing: Predicting Overall Current Performance

The first policy capturing hypothesis predicted that ratings of task-specific performance and interpersonal performance would be significantly related to peer and trainer ratings of overall performance, with ratings of interpersonal performance showing a significantly stronger regression coefficient for peers than trainers. Effort was not expected to have a significant direct effect for either rater group. Results indicated that this hypothesis was partially supported.

Ratings of overall performance were regressed simultaneously on ratings of (1) task-specific performance, (2) interpersonal performance, and (3) effort, for both trainers and peers. With respect to task-specific and interpersonal performance, results for peer ratings showed the predicted significant regression coefficients for both predictors (see Table 10). For trainer ratings, however, the coefficient for task-specific performance was significant, but that for interpersonal performance was not.

As predicted, interpersonal performance ratings had a significantly stronger regression coefficient for peers (Bp=.17, $p<.01$) than for trainers (Bt=.09, ns) (Bp-Bt=.08, $p<.05$). Contrary to predictions, however, effort showed a significant direct effect for both peers and trainers. This suggested that information concerning the amount of effort an individual expends was interpreted by both peers and trainers as directly relevant to the individual's overall performance, not just to the component task-specific and interpersonal performance dimensions.

**Table 10. Comparison of Regression Coefficients for Peer and Trainer Raters**

*1. Overall Current Performance regressed on Component Ratings:*

| | Peers | | Trainers | | $B_p$-$B_t$ |
|---|---|---|---|---|---|
| | *B* | $R^2$ | *B* | $R^2$ | |
| **Task** | .64** | | .70** | | -.06 |
| **Interp** | .17** | | .09 | | .08* |
| **Effort** | .22** | .95 | .17** | .84 | .05 |

*2. Future Performance regressed on Component Ratings:*

| | Peers | | Trainers | | $B_p$-$B_t$ |
|---|---|---|---|---|---|
| | *B* | $R^2$ | *B* | $R^2$ | |
| **Task** | .46** | | .57** | | -.11* |
| **Interp** | .28** | | .16** | | .12** |
| **Effort** | .30** | .93 | .20** | .77 | .10 |

Note: *$p<.05$ and **$p<.01$

Predicting Future Performance

The second policy capturing hypothesis predicted that the regression coefficients for trainers would remain the same using future performance as the criterion, but the regression coefficient for peer effort ratings would evidence a significant direct effect. This precise hypothesis was rendered false when effort ratings were previously shown to be significantly related to current performance for both peers and trainers. It could still be hypothesized, however, that the regression coefficient for effort in the peer model would increase significantly when current performance was replaced with future performance, and the coefficient for effort in the trainer model would not. Results supported this modified hypothesis (see Table 11).

Regression coefficients obtained for task-specific, interpersonal, and effort ratings, using future performance as the dependent variable, were compared with those obtained using current overall performance, for both peers and trainers. As predicted, regression weights for trainers for interpersonal and effort ratings were not significantly different when predicting future and current performance ratings (see Table 11). The coefficient for task-specific performance, however, decreased significantly from Tc=.70 ($p<.01$) for current performance to Tf=.57 ($p<.01$) when predicting future performance (Tf-Tc= -.13, $p<.01$).

When predicting future performance, peer ratings showed the predicted significant increase in the regression weight for effort (Pf-Pc=.08, $p<.05$), as well as an unexpected increase in the coefficient for interpersonal performance (Pf-Pc=.11, $p<.01$), and a large decrease in the coefficient for task-specific performance (Pf-Pc= -.18, $p<.01$). The amount of weight trainers gave to interpersonal performance did increase from B=.09 (*ns*) for current performance to B=.16 ($p<.01$) for future performance; however this difference was not statistically significant.

**Table 11. Difference in Regression Coefficients using Current and Future Performance as Criteria**

*1. Differences in Peer Predictor coefficients for each criterion:*

| | Criteria | | |
|---|---|---|---|
| | **Current Performance** | **Future Performance** | **$B_f$-$B_c$** |
| Regressed on: | $B_c$ | $B_f$ | |
| 1. Task | .64 | .46 | -.18** |
| 2. Interp | .17 | .28 | .11** |
| 3. Effort | .22 | .30 | .08* |

*2. Differences in Trainer Predictor coefficients for each criterion:*

| | Criteria | | |
|---|---|---|---|
| | **Current Performance** | **Future Performance** | **$B_f$-$B_c$** |
| Regressed on: | $B_c$ | $B_f$ | |
| 1. Task | .70 | .57 | -.13* |
| 2. Interp | .09 | .16 | .07 |
| 3. Effort | .17 | .20 | .03 |

Note: *$p<.05$ and **$p<.01$

This indicated that, as predicted, effort significantly increased in importance to peer raters when estimating an individual's future performance on the job; it did not, however, increase in importance to trainers. In addition, interpersonal performance increased significantly in importance

to peer raters when estimating an individual's future performance on the job; while it did increase in importance to trainers, this increase was not statistically significant. Coincident with these increases, the importance of task-specific performance decreased significantly for both peers and trainers.

Comparisons between regression coefficients for the two models when predicting future performance indicated that the regression coefficient for interpersonal performance was significantly greater for peers (Bp=.28, $p<.01$) than for trainers (Bt=.16, $p<.01$) (Bp-Bt=.12, $p<.05$). A similar trend was seen for effort, where the coefficient for peer effort (Bp=.30, $p<.01$) was considerably greater than that for trainer effort (Bt=.20, $p<.01$), although this difference of .10 was not significant. In addition, the regression coefficient for peer task-specific performance (Bp=.46, $p<.01$) was significantly lower than that for trainer task-specific performance (Bt=.57, $p<.01$) (Bp-Bt=.11, $p<.05$).

While 95% of the variance in peer ratings of current performance was explained, 84% of the variance in trainer ratings of current performance was explained. This difference was even more drastic for future performance ratings; while 93% of the variance in peer ratings of future performance was explained, only 77% of the variance in trainer ratings of future performance was explained.

## DISCUSSION

### Causal Models: Predictors of Performance Ratings

#### Task-specific performance.

Results for the first hypothesis indicated that ratings of task-specific performance were predicted by experience and effort. The strong effect of experience appeared to preclude the importance of cognitive ability, and cognitive ability did not directly affect performance. The direct correlation with performance may also have been attenuated by the range restriction on cognitive ability in this sample. For peer raters, some of the variance in task-specific performance was explained *directly* by the student's task-specific self-efficacy. This suggests that an individual's confidence in performing a specific task directly increased his peer-rated task performance, and suggests that peer evaluations may provide more accurate information than trainer ratings, a finding which is consonant with previous research (McCloy, Campbell, & Cudeck, 1994).

Two other possible explanations for the direct effect from task self-efficacy to peer ratings of performance do exist, however. First, it is possible that the direct link from task self-efficacy to performance reflects the fact that an individual's confidence in performing a specific task is verbally communicated to his peers; when peers know that an individual has expressed confidence about his performance, they tend to rate him higher, regardless of demonstrated performance. This explanation would suggest that peer evaluations are actually less accurate because this type of error is introduced into peer ratings. A second possible argument is that the link between task-specific efficacy and task performance is not actually causal, but correlational, and it simply reflects the fact that people are generally good judges of their own abilities or performance potential. These alternate explanations are less plausible, however, because they are not consonant with research from the self-efficacy literature that has shown direct links between task self-efficacy and objective measures of performance (e.g., Ryckman, Robbins, Thornton, & Cantrell, 1982; Wood & Bandura, 1989; Deeter, 1990; Wurtele, 1986; Wood & Locke, 1987).

#### Interpersonal Performance.

Ratings of interpersonal performance were predicted directly by effort for both peers and trainers, and predicted indirectly by experience. For trainers, the effect of experience on interpersonal performance was direct as well as indirect. While previous performance rating models have addressed interpersonal factors by including personality constructs, this model introduced a measure of social intelligence to predict interpersonal efficacy and interpersonal performance. The path from social intelligence to interpersonal performance was not supported; however, the measure was predictive of interpersonal self-efficacy. Given that it is a relatively distal measure, it is not surprising that its effects would be mediated through efficacy.

#### Comparison to previous models.

These results demonstrated that the proximal motivational variables added to this performance model explained considerably more variance in performance ratings than models examined in previous studies. Hunter's (1983) original model, which included only "can-do" predictors,

explained 16% of the variance in ratings based on job knowledge and work sample scores. Borman et al. (1991) added self-report personality characteristics and two proximal performance indicator variables, and increased the amount of variance explained to 31%. Finally, Borman et al. (1994) replaced the self-report personality measures and proximal performance indicator variables with several proximal personality measures (i.e., constructs rated by supervisors and peers), and explained a similar amount of variance (28% in supervisor ratings and 19% in peer ratings). With effort included in the current models as a proximal motivational variable, predictors were able to explain 60-80% of the variance in performance ratings.

Path analyses showed that effort was the primary contributor in explaining both variance in task performance and interpersonal performance, and these high rates of explained variance were expected, given that both the effort rating and performance ratings represented subjective integrations of behavioral observations. The current results have doubled the variance explained by predictors in Borman et al.'s 1994 study by including a proximal measure focused specifically on *motivation*.

One possible criticism of these results is that the magnitude of the relationship between effort and the performance ratings is either inflated by, or completely the result of, method bias. While Borman et al. (1994) minimized method bias by using different raters to contribute the personality and performance ratings, this was impossible in the current analyses because only one trainer was available to make all trainer ratings. Since it was impossible to use separate raters for effort and performance in the trainer model, results reported for the peer model also used the same rater for effort and performance.

In order to address the issue of method bias, however, results were also calculated for the peer model using different peer raters for effort and performance. Path analysis results demonstrated that when different peers were used to make the ratings, the total percentage of variance explained in peer task-specific performance decreased by only 2% to 78%. For interpersonal performance ratings, the total percentage of variance explained remained at 59%. This suggests that, at least for peer raters, using different raters for effort and performance provided nearly identical conclusions.

Causal Models: The Role of Motivational Antecedents

A critical goal of the path models in this research was to add motivational variables that could illustrate causal paths predicting performance. While previous models had only included achievement orientation (Borman et al., 1991; Borman et al., 1994); the current model included self-reported achievement orientation, mastery beliefs, locus of control, and self-efficacy, as well as peer and trainer ratings of effort.

Effort.

Effort reflected the extent to which peers and trainers observed student behaviors related to persistence and determination, and it was the motivational construct that was most proximal to performance in this model. While student effort was expected to be solely predicted by task and interpersonal self-efficacy, effort was actually predicted by experience. While task-specific self-efficacy provided a significant link in the original path models, once the link from experience was added, the link from task-specific self-efficacy became nonsignificant. These results suggested that

students who had previous experience relevant to the course were more likely to be described by peers and trainers as demonstrating high levels of persistence and determination.

Interpersonal self-efficacy failed to show the predicted significant relationship with level of effort. This is especially surprising because the measure did have a high reliability, and its observed relationship with social intelligence suggested that the measure had some validity. The format of the measure was new, however, and may have been somewhat awkward (see Appendix C, items 12 - 14); it asked the student how he believed he would be ranked by his peers with respect to behaviors such as "getting along with others" and "reading people and social situations". It is possible that this item style placed too much emphasis on the role of others in the estimate of one's confidence, instead of the role of the self. Nevertheless, given that the 95% confidence intervals for the coefficient between interpersonal self-efficacy and effort ranged up to .29 for peers and .35 for trainers, future research should continue to examine the utility of this construct, perhaps experimenting with alternate methods of measurement.

Self-efficacy.

Task self-efficacy was hypothesized to be predicted by personality/beliefs, cognitive ability, and experience, and interpersonal self-efficacy to be predicted by personality/beliefs and social intelligence. Results indicated that task self-efficacy was actually predicted by prior experience and cognitive ability scores, and interpersonal self-efficacy was predicted by social intelligence. None of the personality/belief measures were significant predictors of self-efficacy. In these results, while the path from experience confirmed the importance of prior performance in developing self-efficacy beliefs in a new situation, the significant negative path from cognitive ability was the opposite of what was predicted.

One possible reason for this unexpected relationship could be that it is spurious. If this is true, another variable should be present that is correlated with both Wonderlic and efficacy; however, no such variables exist in the model. Task and social self-efficacy were correlated with experience, for example, but the correlation of experience with Wonderlic was not significant; thus, even with experience partialled out of the correlation between Wonderlic and efficacy, the correlation remains significant and negative. While it may be possible that an unmeasured variable caused a spurious relationship in this model, the available evidence suggests that individual differences in cognitive ability were, in fact, negatively related to an individual's prediction of his skills and abilities.

Very limited prior research on self-efficacy has investigated an "over-confidence effect"; that is, the existence of efficacy beliefs that are unreasonable and "overconfident" with respect to an individual's true abilities and skills (e.g., Stone, 1994). Individuals who are less able to accurately monitor their task or interpersonal actions within a situation may be more likely to report unreasonably high confidence for performance. Results from the current research suggested that lower intelligence students were reporting higher self-efficacy; these students were not necessarily performing better, which would suggest that they were overestimating their skills and abilities. In fact, previous research has shown that soldiers with lower cognitive ability scores tended to score lower on the Lennox and Wolfe (1984) measure of self-monitoring (Zaccaro, Zazanis, Diana, & Gilbert, 1995).

A plausible explanation for the negative relationship between cognitive ability and self-efficacy, then, is that lower cognitive ability students were less able to rationally estimate their abilities and skills, and were more likely to overestimate, generating a negative relationship between cognitive ability and self-efficacy. This raises the critical issue of the rational and emotional components of self-efficacy; when subjects respond to self-efficacy items, it is not clear the extent to which their responses are driven by a rational consideration of skills and abilities versus emotionally-based beliefs. While the emotional, motivational dimension of self-efficacy is often the focus of theoretical investigations of the construct, these results highlight a factor potentially affecting the rational dimension of the estimation.

Post-hoc analyses also suggested the importance of context beliefs to each of the self-efficacy constructs. Initial analyses with this variable showed a direct effect of beliefs about the course environment on both task and interpersonal self-efficacies. Patterns of correlations suggested the utility of splitting the construct into a set of context beliefs about the student's teammates and beliefs about the course structure and trainers.

Personality/belief measures.

While efficacy and effort played a role in these models, the role of the motivational personality measures (achievement orientation, mastery beliefs, and locus of control) in the model was nonexistent, despite the fact that the measures displayed reasonable reliabilities and some convergent and divergent validity among the predictors. Previous research has indicated that prior performance is the strongest determinant of self-efficacy (Bandura, 1994), so in the face of strong prior performance information, personality and other variables could become irrelevant. Experience in this research apparently functioned as such a strong determinant of self-efficacy that personality/beliefs were irrelevant in predicting self-efficacy.

Evidence of this effect can be observed by removing individuals who have the strongest prior experience from the sample: Rangers. Rangers had the strongest prior performance information, given that many aspects of the SFQC curriculum had been included in Ranger school. An interesting question, then, was whether removing Rangers from the sample would unveil a relationship between personality/belief constructs and self-efficacy for the students without this directly relevant prior experience. In fact, when Rangers were removed from the sample, mastery beliefs showed a significant zero-order correlation with task-specific self-efficacy ($r=.21$, $p<.05$). The correlations for need for achievement and locus of control, however, remained near zero and nonsignificant.

These results reinforce the theory that prior performance information is the strongest predictor of self-efficacy and suggest that when prior performance information is *not* available an individual's mastery beliefs may guide the development of self-efficacy beliefs, and ultimately, influence performance. Although an internal locus of control was significantly related to high mastery beliefs, the fact that only mastery beliefs, out of the three belief constructs, was related to self-efficacy in the non-Ranger sample suggests that it may be a more relevant construct in modeling the path of motivation to performance ratings.

The fact that the new mastery beliefs measure predicted self-efficacy with Rangers removed from the sample argues for its continued investigation in adult populations. Additional research clearly needs to be done to compare and contrast alternative methods for measuring this construct

and examine the issues of domain specificity and generalized versus self-focused items. In order to investigate these motivational personality variables with respect to performance ratings, future research efforts would need to use a setting in which experience plays a smaller role in the development of self-efficacy. The effect of personality variables on the development of self-efficacy will be maximized when individuals have less prior experience on which to base efficacy judgments.

The Role of Cognitive Ability and its Link to Self-efficacy

Cognitive ability displayed an extremely unusual and unexpected pattern within these models. Previous models had each shown positive indirect effects of ability on performance ratings, although the magnitude of these effects varied from .26 for Hunter (1983), to .06 for Borman et al. (1991), to .06 (peer) and .07 (supervisor) for Borman et al. (1994). Without the intermediary job knowledge and proficiency constructs, it was predicted that cognitive ability would evidence a direct relationship with performance ratings, as well as an indirect effect through self-efficacy. The current model did not show a direct effect of ability on performance ratings, and the indirect effect through self-efficacy was in the *negative* direction. As discussed previously, this unusual finding may be the results of an overconfidence effect for lower cognitive ability students.

The results for task self-efficacy suggested that higher cognitive ability students were perhaps more realistic in their estimate of their skills and abilities. It is possible that this result stems partly from the instructions that were given to students in filling out this measure; they were instructed that we were "interested in how accurately you can estimate your skills." These instructions were used in order to create more realistic variance and to decrease negative skew in the self-efficacy scores. It is possible that, especially in the case of low experience students, the high cognitive ability students were more rational or harsher in estimating their current skill level.

This raises an interesting research question regarding cognitive ability, self-efficacy, and performance. The results clearly demonstrated that students with high levels of experience performed well. For students with low levels of experience, the high cognitive ability students in this research reported lower self-efficacy, and the lower cognitive ability ones reported higher self-efficacy. Both of these constructs (ability and self-efficacy) have demonstrated positive relationships with performance in previous research. An interesting question, then, is which of these constructs is better at predicting performance within the low experience group.

Unfortunately, this question is difficult to answer with this small sample. Selecting the subset of low experience students (a sample of only 57), shows a higher magnitude for the correlation between Wonderlic and peer task performance (r=.22, ns) than between task self-efficacy and peer task performance (r=-.01, ns), although neither correlation is significant. (Correlations for trainer ratings are closer in magnitude, at -.02 and .12 for Wonderlic and self-efficacy, respectively.) The peer rating correlations within the subsample are particularly interesting when they are contrasted with the original full-sample correlation matrix (see Table 4). The correlation between Wonderlic and peer task performance was small (r=-.04,ns), and the correlation between task self-efficacy and peer task performance ratings was larger in magnitude and significant (r=.31, $p<.01$). These results suggest that, without the benefit of experience or high ability, high self-efficacy provided little benefit in improving performance.

While these results are exploratory and inconclusive, they reflect the critical issue of the rational and emotional components of self-efficacy; when subjects respond to self-efficacy items, it is not clear the extent to which their responses are driven by a rational consideration of skills and abilities versus emotionally-based beliefs. Responses from lower ability subjects may place too much emphasis on a purely emotional dimension of self-efficacy.

The Role of Experience

A student's prior military experience had an overwhelmingly strong effect on his task-specific performance in this model, both directly, as well as indirectly, through its effect on effort. While the zero-order correlation between experience and supervisory ratings in the Hunter et al. (1986) model was .33, the zero-order correlation between experience and task performance ratings in the current model was .51 for trainers and .61 for peers. With job knowledge and work sample performance in the Hunter et al. (1986) model, experience evidenced only indirect effects on ratings, through job knowledge and work sample performance. Without these variables, experience in this research showed both a direct effect on performance ratings, as well as an indirect effect through effort. This unpredicted path from experience through effort was a critical omission in the model. This indicated that, not only did prior relevant experiences affect an individual's performance directly, presumably through learned skills (not measured here), it affected performance by raising or lowering an individual's persistence and effort toward completing the tasks.

The strength of experience in the model reduced the importance of personality and beliefs in predicting self-efficacy. This problem stemmed from the fact that, largely because of the presence of Rangers, experience was unexpectedly functioning as a strong prior performance construct in its effect on self-efficacy, as opposed to functioning as a more general "relevant experience" measure. When Rangers were removed from the sample, a different pattern of results emerged; a significant positive relationship was shown between trainability beliefs and task and interpersonal self-efficacies. This suggests that different models were operating to predict performance, based on whether prior performance information was or was not available to subjects. With a larger sample size, the model analyses could be completed separately based on the availability of prior performance information.

Because neither the peer nor trainer raters were necessarily blind to a student's previous experience, it could be argued that the strength of the relationship between experience and performance ratings was created by the rater having knowledge of this previous experience, as opposed to through observation of differences in performance. It is important to consider the extent to which the strong correlations between experience and both peer and trainer ratings were generated by a perception bias as opposed to a true difference in performance. Evidence from two sources suggests that a true relationship does exist between experience and performance, although a small proportion of the relationships may be due to perception bias.

One source was a comparison of the magnitude of the relationships between experience and task and social ratings of performance; for both peer and trainer ratings, experience showed a stronger relationship with ratings of task performance than social performance. For peers, in fact, the zero-order correlation between experience and the task performance rating was double the correlation with interpersonal performance ($r=.61$, $p<.01$ versus $r=.30$, $p<.01$, for task and interpersonal, respectively). The fact that experience was more strongly related to task performance than

interpersonal performance demonstrates divergent validity for the experience-performance relationship.

The second source of information reflecting the validity of experience-performance relationship was the relationship between experience and performance during the assessment center which selects soldiers for this training course. The assessment and selection process includes 21 days of physical, land navigation, and basic field assessments, during which soldiers wear no identifying information and are referred to only by a randomly assigned roster number, worn on an arm band. Lappin (1995) reported that more experienced soldiers (i.e., higher ranks, combat arms backgrounds, and Rangers) performed better during the selection process, and, in fact, the more experienced the soldier, the higher his performance across different assessment tasks. The combination of these two pieces of information argues that the relationship between experience and ratings is based on a real effect, not simply an artifact of perception.

One final issue should be mentioned regarding experience in these models. For trainers, in addition to directly affecting efficacy, effort, and task-specific performance, experience also directly influenced interpersonal performance. For peers, the entire effect of experience on interpersonal performance was mediated by effort. Course personnel suggested it was likely that high levels of prior experience made some soldiers more similar to trainers, causing them to interact more appropriately and confidently with trainers. This, in turn, would have created a direct link between experience and trainer ratings of interpersonal performance, but not peer ratings of this construct.

Utility of Task/Interpersonal Performance Split

While previous models did not separate the task-specific and interpersonal dimensions of the performance ratings (e.g., Hunter, 1983; Borman et al., 1991; Borman et al., 1994), the current model separated these ratings and found considerable differences between the antecedents specified for each of these dimensions, though not necessarily for the predicted paths. In the peer model, while interpersonal performance was directly predicted only by effort, task-specific performance was directly predicted by experience and task-specific efficacy in addition to effort. In the trainer model, interpersonal performance and task-specific performance were both predicted by experience and effort, yet the coefficient for the link from experience to task performance was nearly double the coefficient from experience to interpersonal performance. These divergent results clearly suggest the utility of separating task and interpersonal dimensions in models delineating antecedents to performance ratings, especially if there is a large interpersonal component to the job.

Peer-Trainer Differences

It was predicted that the proposed model would fit the data better using peer raters than trainer raters. Although these results showed that motivation was an important antecedent to performance ratings, they did *not* demonstrate a clear difference in the importance of the motivational or social antecedents for the causal models predicting peer and trainer performance ratings. It had been predicted that task self-efficacy would explain significantly more variance in peer observations of effort than trainer observations of effort, and that social intelligence would have a significantly

stronger direct effect on peer ratings of interpersonal performance than trainer ratings of interpersonal performance.

Although these planned contrasts did not provide support to the basic proposition that peer evaluations contain more variance than trainer evaluations predictable by individual motivation, there were several aspects of the results that did offer some support for this: (1) peers showed more discriminant validity for task self-efficacy predicting task and interpersonal performance ratings than did trainer ratings, (2) confidence intervals for the peer model were narrower and less negative, and (3) the peer model evidenced an unexpected direct link from task self-efficacy to task-specific performance, and explained more variance in task-specific performance than the trainer model.

The failure of these data to demonstrate stronger differences could have a number of causes: (a) few or no differences generally occur in the antecedents of peer and trainer ratings of performance, (b) this research used an inappropriate setting or sample, or (c) this research used inadequate or inappropriate antecedent measures. The following sections will address each of these possibilities.

Determining the existence of true differences.

Despite the fact that the proposed hypotheses failed to show differences between antecedents for the two models, results did show that task self-efficacy had a significant direct effect on peer ratings of task-specific performance that did not exist in the trainer model. This added direct effect resulted in a total effect for self-efficacy on task performance of .11 for peers, with no significant effect for trainers. Although a post-hoc finding, this result provides an impetus to maintain the possibility that there are differences in the antecedents to performance ratings for peers and trainers.

Setting or sample characteristics.

With respect to the sample, while the sample size is adequate for the minimum ratio of subjects to predictors, it is relatively small; a larger sample size may produce more stable results. In addition, as discussed previously, having Rangers in the sample appears to be negating the relevance of most antecedents except experience. In order to identify differences in the relevance of motivational personality characteristics, it would be useful to have a sample that is more homogeneous and less experienced.

Finally, with respect to the setting, interviews with the trainers suggested that the situation used for this research may not have been optimal for identifying these differences between peers and trainers, since there may be more opportunity for prolonged student-trainer interactions in this setting than in the typical training setting. Specifically, training was conducted intensely over 25 days, during many different times of the day and many different conditions. While the trainer clearly maintained a different relationship with students than students had with each other, it is highly possible that the nature of this situation was detrimental to exposing differences between the perspectives of peers and trainers.

In addition while trainers did not have access to peer ratings prior to completing their own ratings, interview information suggested that some trainers may have had knowledge of the peers' perceptions of a student prior to completing ratings on his students. This exposure may have served

to diminish the differences in information between the two groups of raters. The existence of these situational factors suggests that it would be worthwhile to re-examine these models in a different training setting; one in which trainer-student interaction is more typical.

Inadequate measures.

The final issue, the adequacy of the measures, is somewhat difficult to determine. Most of the measures showed acceptable reliabilities and at least some construct validity. Achievement orientation was one exception to this, however, and failed to be related to any other construct in the model. This would suggest it may not be useful to maintain the construct, using this measurement, in the model.

## Policy-Capturing Models

These analyses were designed to examine the premise that peer raters would include more information than trainers about interpersonal performance and effort in their ratings of an individual's overall current and future performance; this premise was supported. While both peers and trainers included information about task-specific performance and effort in their overall assessment of an individual's current performance, only peers included unique information about interpersonal performance in this assessment. In addition, when predicting future performance, while both peers and trainers significantly decreased the amount of unique information they included regarding task-specific performance, only peers *significantly increased* the amount of information they included concerning interpersonal performance, and *significantly increased* the amount of information they included concerning effort.

It is this aspect of the structure of peer assessments that may account for their historical strength as assessment tools, especially in predicting future performance (Downey & Duffy, 1978; Kraut, 1975; Wherry & Fryer, 1949; Williams & Leavitt, 1947). Whether this is, in fact, the aspect of peer evaluations that can account for their predictive strength is something that cannot be directly answered by these data, but would require a longitudinal analysis that followed these individuals to their actual performance on-the-job.

Results from the path analyses had indicated that observed effort behaviors affected ratings of an individual's performance in task-specific and interpersonal domains. The policy-capturing regressions have demonstrated that, not only does effort influence domain-specific performance ratings, it also exerts a direct effect on the assessment of an individual's overall performance; therefore, a student's overall performance rating is not simply a measure of his task-specific performance in the training program.

For peers, overall current and future performance was determined by unique information concerning (1) his task-specific performance, (2) the effort and persistence behaviors he demonstrates, and (3) his interpersonal performance. For trainers, overall current performance was determined by unique information concerning (1) his task-specific performance, and (2) the effort and persistence behaviors he demonstrates; predicted future performance is determined by these factors as well as (3) his interpersonal performance. With these three factors, the unique and overlapping variance explained *nearly all* of the variance in peer assessments of overall performance

(95%), and considerable variance in trainer ratings (84%). For peers, the importance of interpersonal skills and effort increased significantly when they were asked to extrapolate and estimate an individual's future performance.

The fact that a lower amount of variance was explained for trainer ratings of current performance (84%) than peer ratings (95%), and a considerably lower amount of variance was explained in future performance for trainers (77%) than peers (93%), could either suggest that trainer ratings contained more error variance, or that trainers have insights into other dimensions that are relevant to current or on-the-job performance, but were not measured here. Other possible influences might include information that trainers may have had about written test scores or the influence of a student's physical appearance.

One argument that could be made against the validity of these results is that the higher coefficients were obtained for the peer ratings due to the higher reliability of those measures. While this is possible, a more in depth consideration of this issue strongly suggests that reliability differences are *not* the sole contributing factor in creating these results. First, although the trainer ratings were assumed to be less reliable than the peer ratings, therefore potentially attenuating results using trainer ratings, the distributions for the peer ratings were significantly skewed, while those for trainer ratings were normal. Thus, results for peer ratings could have been somewhat attenuated due to significant skew, while results for trainer ratings could have been somewhat attenuated due to lower reliability.

Further, there is a clear contrast in the information given by the results for task-specific performance and that for effort and interpersonal performance; although the regression coefficient for interpersonal performance was significantly stronger for peers than trainers, the coefficient for *task*-specific performance, when predicting future performance, was significantly greater for *trainers* than *peers*. If, in fact, reliability in trainer ratings was attenuating their regression coefficients, it seems unlikely that this attenuation would occur only for interpersonal performance and effort, and not for task-specific ratings. These arguments suggest that unreliability was not the primary factor in creating differential regression coefficients for the peer and trainer rating models.

Despite the fact that the reliabilities of the ratings do not appear to play the primary role in creating the differential results for the peer and trainer regressions, the fact that peer and trainer ratings have such largely different psychometric properties warrants investigation in future research. Psychometric analyses showed that the means for all peer ratings were significantly greater than those for the trainer ratings, the variances for the trainer ratings were significantly greater than those for peers, and peer ratings had significant skew while trainer ratings did not. Previous research presents reliability estimates for peer and supervisor ratings but does not address the issue of distributional properties (e.g., Borman et al., 1994; Oppler et al, 1994; McCloy et al, 1994), and it is not clear whether this finding is unique to this setting, or is generalizable. Reasons for these differences in distribution should be investigated, as well as the potential consequences of these differences for the use of peer and trainer ratings as criteria.

## Integrative Discussion

The goals of this research were to demonstrate that ratee motivation and interpersonal skills affect performance ratings, that we can delineate causal paths to describe the effects of motivation and interpersonal factors on performance ratings, and that there are differences in peer and trainer assessments of performance that are rooted, at least in part, in their assessments of motivation and interpersonal skills.

### Motivation and performance.

Results from the path analyses and the policy-capturing models emphasized the importance of motivation in modeling performance ratings. The policy-capturing models demonstrated the significant unique contributions of effort in predicting an individual's overall performance as well as predicting his future performance.

The causal models emphasized ratee motivation as a critical antecedent to the performance ratings a student received, and began to delineate the path taken by motivational constructs in affecting performance ratings. Analyses uncovered information regarding motivational antecedents that were more proximal to performance (i.e., effort and self-efficacy), but only weak information regarding distal antecedents (i.e., personality/beliefs). While previous models had not delineated a path relating motivational antecedents to performance ratings, the models tested in this research were able to explain a significant amount of variance in each of the endogenous motivational antecedents to performance: 13% of the variance in peer ratings of effort, 10% in trainer ratings of effort, and 18% and 38%, respectively, in task-specific and interpersonal self-efficacy. Results described basic paths leading from experience to self-efficacy to performance ratings, and experience to effort to performance ratings, with multiple direct and indirect effects.

### Self-efficacy.

Task self-efficacy showed a tentative presence in these models. Intitial models indicated that higher task self-efficacy during training resulted in higher levels of effort, and consequently, higher levels of task-specific performance; that is, students who had more confidence that they could successfully perform the required tasks, exerted more effort and were rated by peers and trainers as performing more successfully. Overlap in prediction between task self-efficacy and experience resulted in a final model that suggested that the amount of effort an individual exerted was completely determined by their level of experience.

For the peer model, however, task self-efficacy directly predicted peer ratings of task-specific performance. For the performance modeling literature, these results argue that self-efficacy is a critical motivational antecedent to performance, and should be included in future performance modeling research. For the self-efficacy literature, these results demonstrate that the self-efficacy construct is not only relevant to objective measures of performance, but also to *ratings* of performance. Previous research had demonstrated the relevance of self-efficacy to the prediction of objective performance measures (e.g., Ryckman et al., 1982; Wood & Bandura, 1989; Deeter, 1990; Wurtele, 1986); these path models extend these findings, showing that self-efficacy is also predictive of performance ratings.

With respect to the antecedents of self-efficacy, mastery beliefs and especially the environmental responsiveness construct showed considerable promise. For mastery beliefs, relevance as an antecedent appears strongly dependent upon having a situation in which prior performance does not dominate the development of self-efficacy. This construct needs considerably more research with adult populations, since research in this domain has been completely dominated by child subject populations. Dweck and her colleagues demonstrated that these beliefs were changeable in children (see Chiu et al., 1994 for a review); whether mastery beliefs are malleable in adult populations is untested. If these beliefs *are* malleable and they are related to self-efficacy and task performance, understanding *how* to change these beliefs would clearly be important, both to training and organizational settings.

Interpersonal performance.

Results from the path analysis left the antecedents to interpersonal performance primarily unexplained, with the exception of the relevance of effort. Given the importance of interpersonal performance to overall performance shown in the policy-capturing models, it is critical that we investigate potential precursors to interpersonal performance, further examining social intelligence, team environment, and social self-efficacy to strengthen their measurement and identify additional relevant factors.

While the path analysis did not demonstrate the predicted differences in antecedents for peer and trainer ratings of interpersonal performance, the policy-capturing models clearly demonstrated a difference in the importance of interpersonal performance to peer and trainer ratings of overall current and future performance. In addition, these results emphasized that there are differences in either the *ability* of peers and trainers to assess interpersonal factors, or in the *judgment* of peers and trainers of the *importance* of interpersonal factors to assessments of current and future performance.

While Mumford (1983) suggested that the strength of peer evaluations lies in the *superior accuracy* of the observations peers make regarding *task* performance, these results suggested that the strength of peer evaluations lies in the nature of the interpersonal and/or personality content of the peer evaluations. As opposed to focusing on the accuracy of the peer versus trainer assessments, that is, these results suggested the criticality of the content and scope of the peer evaluations. Peers insights into interpersonal skills and personal motivation appear to provide a more complete and complex perspective than that held by trainers or supervisors.

While an individual's knowledge and skills may change considerably over time, and the knowledge and skill requirements of the positions he/she holds may also change considerably over time, the personality and motivational requirements of the positions may change considerably less. Given that peer evaluations reflect interpersonal and motivational characteristics, this would explain their strength in predicting performance over time and across performance situations. Of course, to address this empirically it would be necessary to follow these students through performance in the field and obtain on-the-job performance information to compare the predictiveness of peer and trainer assessments.

Situational Constraints.

There are several constraints on the generalizability of these results that must be mentioned. Most critically, the sample was clearly not randomly selected. Subjects were all military personnel and they were all male. While this does not preclude these results from generalizing to other situations, it should be carefully considered when applying these results to settings other than military, and to settings including female subjects.

In addition, these analyses identified external factors that were significant threats to the current model, and must be carefully considered in designing future research in this domain. One is the criticality of considering the amount of variance that exists in the prior experience and cognitive abilities of a potential subject population; if there is a large amount of variance in experience within a given population, experience is likely to be the most important predictor of performance in that setting, and completely overshadow the importance of ability or personality measures.

Recommendations for research.

While it had been expected that cognitive ability could fulfill the basic role of can-do antecedents in these models, the failure of cognitive ability to predict performance suggests that it is necessary to examine the distal-to-proximal paths for both "can-do" and "will-do" variables within a single sample. It is unfortunate that job knowledge and task proficiency predictors were not available for inclusion in this model; while previous models did not clearly delineate the path of motivational variables from distal to proximal predictors to performance, the model examined here was unable to include the established path for knowledge and skill variables. Including both sets of variables will allow an examination of the interrelationships among the can-do and will-do antecedents. Clearly future research should attempt to examine a model that includes fully delineated paths to task performance, interpersonal performance, and effort.

It also appears particularly critical that we investigate the relationship between cognitive ability and self-efficacy to determine whether a true negative relationship exists between these two constructs; we must determine whether this relationship is grounded in the fact that lower cognitive ability persons are less able to accurately monitor and report their skills and abilities. Investigating the emotional and rational dimensions of self-efficacy may be particularly useful in enlightening the issue of an overconfidence effect.

REFERENCES

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1994). Regulative Function of perceived self-efficacy. In M.G. Rumsey, C.B. Walker, & J.H. Harris (Eds.), *Personnel Selection & Classification.* Hillsdale, N.J.: Erlbaum.

Barrick, M.R., Mount, M.K., & Strauss, J.P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. *Journal of Applied Psychology, 78,* 715-722.

Borman, W.C. (1974). The rating of individuals in organizations: An alternate approach. *Organizational Behavior and Human Performance, 12,* 105-124.

Borman, W.C. (1991). Job behavior, performance, and effectiveness. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 271-326). Palo Alto, CA: Consulting Psychologists Press.

Borman, W.C., White, L.A., & Dorsey, D.W. (1994). *Effects of ratee task performance and interpersonal factors on supervisor and peer performance ratings.* Paper presented at the 9th Annual meeting of the Society of Industrial and Organizational Psychology, Nashville, TN.

Borman, W.C., White, L.A., Pulakos, E.D., and Oppler, S.H. (1991). Models of supervisor job performance ratings. *Journal of Applied Psychology, 76,* 863-872.

Brooks, J.E. (1997). Special Forces recruiting and manpower planning. In J.E. Brooks & M.M. Zazanis (Eds.), *Enhancing Special Forces* (pp. 7-18). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Campbell, J.P. & Zook, L.M. (1991). *Improving the Selection, Classification, and Utilization of Army Enlisted Personnel: Final Report on Project A.* (Research Report 1597). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A242 921)

Chiu, C., Hong, Y., & Dweck, C.S. (1994). Toward an integrative model of personality and intelligence; a general framework and some preliminary steps. In R.J. Sternberg and P. Ruzgis (Eds.), *Personality and intelligence* (pp.104-134). Cambridge, NY: Cambridge University Press.

Deeter, T.E. (1990). Re-modeling expectancy and value in physical activity. *Journal of Sport and Exercise Psychology, 12,* 86-91.

Downey, R.G. & Duffy, P.J. (1978). *Review of peer evaluation research.* (Technical Paper 342). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Dweck, C.S. (1986). Motivational processes affecting learning. *American Psychologist, 11,* 1040-1048.

Ford, M.E. (1992) *Motivating Humans*. Newbury Park, CA: Sage Publications, Inc.

Gilbert, J.A., Connelly, M.S., Mumford, M.D. & Zaccaro, S.J. (1992). *A background data measure of social intelligence*. Fairfax, VA: Center for Behavioral and Cognitive Studies, George Mason University.

Gough, H.G. & Heilbrun, A.B. (1983). *The Adjective Check List Manual* (rev. ed.). Palo Alto, CA: Consulting Psychologists Press, Inc.

Guion, R.M. (1983). Comments on Hunter. In F.Landy, S. Zedeck, and J. Cleveland (Eds.), *Performance measurement and theory*. Hillsdale, NJ: Erlbaum.

Hough, L.M. (1986). *Literature Review: Utility of Temperament, Biodata, and Interest Assessment for Predicting Job Performance*. (ARI RN 88-02). Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences.

Hunter, J.E. (1983). A causal analysis of cognitive ability, job knowledge, job performance, and supervisory ratings. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory* (pp. 257-266). Hillsdale, NJ: Erlbaum.

Jackson, D.N. (1984). *Personality Research Form Manual*. Port Huron, MI: Research Psychologists Press, Inc.

Jonassen, D.H. & Grabowski, B.L. (1993). *Handbook of Individual Differences, Learning, and Instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kane, J.S. & Lawler, E.E., III (1978). Methods of peer assessment. *Psychological Bulletin, 85*, 555-586.

Kanfer, R. (1991). Motivation theory and industrial and organizational psychology. In M.D. Dunnette & L.M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 75-170). Palo Alto, CA: Consulting Psychologists Press.

King, L.M., Hunter, J.E., and Schmidt, F.L. (1980). Halo in a multidimensional forced choice performance evaluation scale. *Journal of Applied Psychology, 65*, 507-516.

Klimoski, R.J. & London, M. (1974). Role of the rater in performance appraisal. *Journal of Applied Psychology, 59*(4), 445-451.

Knapp, D.J. & Campbell, J.P. (1993). *Building a joint-service classification research roadmap: Criterion-related issues*. (AFMC AL/HR-TP-1993-0028). Brooks Air Force Base, TX: Human Resources Directorate Manpower and Personnel Research Division.

Kraut, A.I. (1975). Prediction of managerial success by peer and training-staff ratings. *Journal of Applied Psychology, 60*, 14-19.

Lappin, M.S. (1995). *The CMF 18 Assessment/Training Pipeline: Attrition and Success Factors*. Briefing presented to the CMF 18 Sustainment Group, Ft. Bragg, NC.

Lennox, R. D., & Wolfe, R. N. (1984). Revision of the self-monitoring scale. *Journal of Personality and Social Psychology, 46*, 1349-1364.

Lewin, A.Y. & Zwany, A. (1976). Peer nominations: A model, literature critique and a paradigm for research. *Personnel Psychology, 29*, 423-447.

McCloy, R.A., Campbell, J.P., & Cudeck, R. (1994). A confirmatory test of a model of performance determinants. *Journal of Applied Psychology, 79*, 493-505.

Mumford, M. D. (1983). Social comparison theory and the evaluation of peer evaluations: A review and some applied implications. *Personnel Psychology, 36(4),* 867-881.

Murphy, K.R. & Cleveland, J.N. (1991). *Performance Appraisal.* Boston, MA: Allyn and Bacon.

Oppler, S.H., Peterson, N.G., & McCloy, R.A. (1994). *A comparison of peer supervisor ratings as criteria for the validation of predictors*. Paper presented at the 9th Annual meeting of the Society of Industrial and Organizational Psychology, Nashville, TN.

Oppler, S.H., Pulakos, E.D., & Borman, W.C. (1992). *Comparing the influences of ratee characteristics on peer and supervisor ratings*. Paper presented at the annual convention of the Society for Industrial and Organizational Psychology, Montreal, Canada.

Reilly, R.R. & Chao, C.T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology, 35*, 1-62.

Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, (Whole No. 609).

Ryckman, R.M., Robbins, M.A., Thornton, B., & Cantrell, P. (1982). Development and validation of a physical self-efficacy scale. *Journal of Personality and Social Psychology, 42*, 891-900.

Stone, Dan N. (1994). Overconfidence in initial self-efficacy judgments: Effects on decision processes and performance. *Organizational Behavior and Human Decision Processes, 59,* 452-474.

Wherry, R.J., & Fryer, D.H. (1949). Buddy ratings: Popularity contest or leadership criteria? *Personnel Psychology, 2*, 147-159.

Williams, S.B., & Leavitt, H.J. (1947). Group opinion as a predictor of military leadership. *Journal of Consulting Psychology, 11*, 283-291.

Wonderlic, C.F. (1988). *Wonderlic Personnel Test.* Northfield, IL: E.F. Wonderlic Personnel Test, Inc.

Wood, R.E. & Bandura, A. (1989). Impact of conceptions of ability on self-regulatory mechanisms and complex decision making. *Journal of Personality and Social Psychology, 56*, 407-415.

Wood, R.E. & Locke, E.A. (1987). The relation of self-efficacy and grade goals to academic performance. *Educational and Psychological Measurement, 47*, 1013-1024.

Wurtele, S.K. (1986). Self-efficacy and athletic performance: A review. *Journal of Social and Clinical Psychology, 4*(3), 290-301.

Zaccaro, S.J., Zazanis, M.M., Diana, M., & Gilbert, J.A. (1995). *Investigating a background data measure of social intelligence* (Technical Report 1024). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A298 832)

Zaccaro, S.J., Zazanis, M.M., Diana, M., & Greathouse, C.G. (1994). *The antecedents of collective efficacy over a team's lifespan*. Paper presented at the 9th annual meeting of the Society of Industrial and Organizational Psychology, Nashville, TN.

Zazanis, M.M., Diana, M., Lappin, M.S., & Schweisthal, D.S. (in preparation). *Special Forces Qualification Course: Background profiles of the successful student (Fiscal year 1992 to fiscal year 1994).* (Research Report, in preparation). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

APPENDIX A
Subsample Comparisons on Existing Variables

**Summary of Subsample Comparisons**

Subjects who had intelligence, personality, and belief data available (N=151) were compared with those who did not (N=88), with respect to the variables that they had in common: Army component, Army grade, peer ratings, and trainer ratings. Results from Table 1 show that the two groups had very similar percentages of students from each of the Army components and grades. Next, means for the peer and trainer ratings were compared for the two subgroups. Results showed no significant differences in the means for any of the peer or trainer ratings.

Finally, the intercorrelation matrices for the ratings were compared for the two subgroups (see Tables 2 - 7). Only 2 of the 45 intercorrelations were significantly different for the two subgroups: the correlation between peer overall and peer social, $Z=2.02$, $p<.05$, and the correlation between peer future and peer social, $Z=2.07$, $p<.05$. Given that only 2 of 45 intercorrelations were significantly different, it is highly likely that this is a chance finding. When a more conservative significance level is used to compensate for the large number of post-hoc tests conducted (e.g. $p<.01$) , these values are not significant. Overall these comparisons indicated no notable differences between the subsample that has the additional data, and the subsample that does not.

**Table 1**

**Component and grade breakdowns for each subsample**

| | N=151 | N=88 |
|---|---|---|
| *Army component* | | |
| Active Duty | 70.2% | 69.3% |
| Reserves | 29.8% | 30.7% |
| *Army grade* | | |
| E-4 | 27.2% | 22.7% |
| E-5/E-6 | 72.1% | 77.3% |

**Table 2**

**Peer Intercorrelations for Subjects with Incomplete Data**

| | PEFFORT | PSOCIAL | PTASK | POVERALL | PFUTURE |
|---|---|---|---|---|---|
| PEFFORT | 1.0000 | .8473** | .8150** | .8881** | .9093** |
| PSOCIAL | | 1.0000 | .7879** | .8810** | .9017** |
| PTASK | | | 1.0000 | .9467** | .9095** |
| POVERALL | | | | 1.0000 | .9723** |
| PFUTURE | | | | | 1.0000 |

** - Signif. LE .01 (2-tailed)

**Table 3**

**Peer Intercorrelations for Subjects with Complete Data**

| | PEFFORT | PSOCIAL | PTASK | POVERALL | PFUTURE |
|---|---|---|---|---|---|
| PEFFORT | 1.0000 | .7880** | .8278** | .8917** | .8983** |
| PSOCIAL | | 1.0000 | .7224** | .7983** | .8304** |
| PTASK | | | 1.0000 | .9570** | .9203** |
| POVERALL | | | | 1.0000 | .9598** |
| PFUTURE | | | | | 1.0000 |

** - Signif. LE .01 (2-tailed)

**Table 4**

**Trainer Intercorrelations for Subjects with Incomplete Data**

| | CEFFORT | CSOCIAL | CTASK | COVERALL | CFUTURE |
|---|---|---|---|---|---|
| CEFFORT | 1.0000 | .7604** | .8117** | .8364** | .8088** |
| CSOCIAL | | 1.0000 | .7287** | .7341** | .7507** |
| CTASK | | | 1.0000 | .9002** | .8820** |
| COVERALL | | | | 1.0000 | .9151** |
| CFUTURE | | | | | 1.0000 |

** - Signif. LE .01 (2-tailed)

**Table 5**

**Trainer Intercorrelations for Subjects with Complete Data**

| | CEFFORT | CSOCIAL | CTASK | COVERALL | CFUTURE |
|---|---|---|---|---|---|
| CEFFORT | 1.0000 | .7671** | .8002** | .7758** | .7665** |
| CSOCIAL | | 1.0000 | .8085** | .7785** | .7641** |
| CTASK | | | 1.0000 | .9073** | .8510** |
| COVERALL | | | | 1.0000 | .8777** |
| CFUTURE | | | | | 1.0000 |

**- Signif. LE .01 (2-tailed)

**Table 6**

**Peer-Trainer Correlations for Subjects with Incomplete Data**

| | CEFFORT | CSOCIAL | CTASK | COVERALL | CFUTURE |
|---|---|---|---|---|---|
| PEFFORT | .6211** | .6388** | .6512** | .6226** | .6634** |
| PSOCIAL | .6280** | .7130** | .6256** | .6454** | .6969** |
| PTASK | .6692** | .6596** | .7822** | .7821** | .7789** |
| POVERALL | .7235** | .7251** | .7916** | .7950** | .8069** |
| PFUTURE | .6961** | .7327** | .7528** | .7473** | .7848** |

** - Signif. LE .01 (2-tailed)

**Table 7**

**Peer-Trainer Correlations for Subjects with Complete Data**

| | CEFFORT | CSOCIAL | CTASK | COVERALL | CFUTURE |
|---|---|---|---|---|---|
| PEFFORT | .6336** | .6366** | .6921** | .6799** | .6804** |
| PSOCIAL | .5354** | .6050** | .5945** | .6030** | .6343** |
| PTASK | .6646** | .7379** | .8439** | .8201** | .7609** |
| POVERALL | .6682** | .7541** | .8085** | .8037** | .7745** |
| PFUTURE | .6744** | .7474** | .7928** | .7852** | .7700** |

** - Signif. LE .01 (2-tailed)

APPENDIX B
Survey of Mastery Beliefs

These belief scales were developed to measure the extent to which soldiers believe that the abilities required for Special Forces training are innate or are trainable or attainable. Five domains that have been identified as crucial to performance in Special Forces selection and training are included: physical, leadership, tactical/operational, academic/technical, and interpersonal.

**Physical**

Agree/Disagree:

1 8. P1. If they trained for it, most soldiers could run about a 5-minute mile.

2 3. P2. Someone could probably never learn to do an obstacle course well if he didn't have natural coordination for it.

1 16.P3. If someone wanted to, they could push their body past its natural limits and do twice as many situps as they usually do.

1 22.P4. If someone did 55 pushups for a PT test one month, they could probably do 80 the next month if they trained for it.

**Leadership**

2 25.L1. Usually it's not possible to teach someone how to influence others - if you can't naturally do it, you won't be able to learn.

1 7. L2. If someone wanted to be a leader, he could train to become a good leader even if he wasn't naturally good at it.

1 18.L3. The Army should spend more time and money on training people to be good leaders than on selecting good leaders from the start.

2 1. L4. Once someone develops a leadership style, it's nearly impossible for them to change it.

**Tactical - Field**

2 2. LN1. A person with low intelligence wouldn't be able to learn land navigation skills.

1 26.LN2. If they trained for it, most soldiers could successfully navigate without a compass or flashlight.

2 21.SP1. If someone is not very good at picturing things "in his head", he wouldn't be successful at learning how to read maps.

**Tactical, cont.**

1 12.SP3. Being good at interpreting a contour/terrain map is mostly due to training, not natural ability.

2 17.SP4. If you weren't born with a good sense of direction, you probably wouldn't be able to learn land navigation.

2 29.T1. Field operations are one part of the military job that requires mostly natural ability.

2 5. T2. If someone doesn't naturally have common sense, he couldn't really be trained to perform well in military tactics.

2 9. T3. People that aren't naturally good at reacting to combat situations would never be able to learn enough to make them perform well in combat.

**Academic - Cognitive Ability (Technical/Book Learning)**

2 27.A1. A person with low intelligence would not be able to learn technical knowledge about his field.

2 13.A2. If someone's not good at memorizing things, he wouldn't be able to pass a class that required a lot of memorization.

2 4. A3. If you're not good at learning from books you wouldn't be able to do well in a classroom situation.

1 10.A4. People that failed math in high school could have gotten "A's" if they had studied harder.

2 20.A5. People that are really good teachers are naturals at it - it's not something you can really learn.

**Interpersonal**

2 14.SO1. If someone doesn't like you when you first meet, there is little you can do to change their mind about you.

2 6.SO2. Some people just seem to say the wrong thing at the wrong time - they would never be able to change, even if they wanted to.

2 24.SO3. Once they get to be adults, most people wouldn't be able to change how they deal with people very much from one situation to the next.

**Interpersonal, cont.**

2 19.SO4. People can't really be trained to know how to interact with other people.

1 28.SO5. Most people could learn how to fit in to a really different culture if they kept working at it.

**Extras**

23.F1. Learning basic math is mostly a matter of training, but learning algebra and geometry also requires a lot of natural ability.
11.F2. Your physical abilities do not limit how well you can perform land navigation.
15.F3. Some people seem to be naturally good at almost everything.

# APPENDIX C

## Survey of Self-Efficacy

Enter your roster number here:

Roster

Directions:

We are investigating students' assessments of their current skill levels. The following questions ask you how confident you are that you could perform tasks that vary in difficulty.

When you answer the questions, answer as accurately as possible, given what you know about your current skills.

This form will be read by a machine, so be sure to fill in your answers carefully and completely.

| | Confident | | | |
|---|---|---|---|---|
| | Not at all Confident | Slightly Confident | Very Confident | Completely Confident |
| 1. If you were in wooded terrain carrying a 50 lb. rucksack, how confident are you that you have the ability to: | | | | |
| a. Find 4 points over a 6 km route in 4 hours | | | | |
| b. Find 4 points over an 8 km route in 4 hours | | | | |
| c. Find 4 points over a 9 km route in 4 hours | | | | |
| d. Find 4 points over a 10 km route in 4 hours | | | | |
| e. Find 2 of 4 points over a 20 km route in 8 hours | | | | |
| f. Find 3 of 4 points over a 20 km route in 8 hours | | | | |
| g. Find 4 of 4 points over a 20 km route in 8 hours | | | | |

| | | |
|---|---|---|
| 2. How confident are you that you have the ability to break down an M16A2 rifle: | Not at all Confident | Completely Confident |
| a. within 2 minutes | | |
| b. within 1 minute | | |
| c. within 30 seconds | | |
| 3. How confident are you that you have the ability to break down an M60 machine gun: | Not at all Confident | Completely Confident |
| a. within 4 minutes | | |
| b. within 3 minutes | | |
| c. within 2 minutes | | |
| 4. How confident are you that you have the ability to correctly emplace a Claymore mine during the daytime: | Not at all Confident | Completely Confident |
| a. in 3 minutes | | |
| b. in 2 minutes | | |
| c. in 1 minute | | |



Form Number 75020-5-72

| | Confident | | | |
|---|---|---|---|---|
| | Not at all Confident | Slightly Confident | Very Confident | Completely Confident |
| 5. How confident are you that you have the ability to correctly emplace a Claymore mine at night:<br>a. in 3 minutes<br>b. in 2 minutes<br>c. in 1 minute | | | | |
| 6. How confident are you that within 1 day you could teach a squad of untrained guerrillas to:<br>a. Move as members of a squad<br>b. Move through a linear danger area<br>c. Move through a large open danger area | Not at all Confident | | | Completely Confident |
| 7. How confident are you that you could plan for and conduct a withdrawal from a patrol base while under fire and under limited visibility. | Not at all Confident | | | Completely Confident |
| 8. How confident are you that you have the ability to plan and issue an OPORD for a squad ambush within:<br>a. 5 hours while maintaining 50% security<br>b. 4 hours while maintaining 50% security<br>c. 3 hours while maintaining 50% security<br>d. 2 hours while maintaining 50% security | Not at all Confident | | | Completely Confident |
| 9. You are the patrol leader for a squad night ambush and you just initiated an ambush on a 2 vehicle convoy. There are 4 enemy KIA and 1 enemy WIA. How confident are you that you have the knowledge and skills required to:<br>a. Clear and withdraw to an ORP within 20 minutes<br>b. Clear and withdraw to an ORP within 10 minutes<br>c. Clear and withdraw to an ORP within 5 minutes | Not at all Confident | | | Completely Confident |
| 10. How confident are you that you could control a squad that is:<br>a. Receiving indirect fire<br>b. In chance contact<br>c. In a near ambush | Not at all Confident | | | Completely Confident |

11. If your peers were asked to rank students on their tactical proficiency, how confident are you that you would be ranked:

Confident

| Not at all Confident | Slightly Confident | Very Confident | Completely Confident |
|---|---|---|---|

a. In the top 50%
b. In the top 30%
c. In the top 10%
d. In the top 1%

12. If your peers were asked to rank students on their ability to get along with other people, how confident are you that you would be ranked:

Not at all Confident — Completely Confident

a. In the top 50%
b. In the top 30%
c. In the top 10%
d. In the top 1%

13. If your peers were asked to rank students on their ability to read people and social situations, how confident are you that you would be ranked:

Not at all Confident — Completely Confident

a. In the top 50%
b. In the top 30%
c. In the top 10%
d. In the top 1%

14. If your peers were asked to rank students on their ability to act appropriately in different situations, how confident are you that you would be ranked:

Not at all Confident — Completely Confident

a. In the top 50%
b. In the top 30%
c. In the top 10%
d. In the top 1%



Form Number 75020-5-72

APPENDIX D

Rating Form

PEER EVALUATIONS

We are asking you to evaluate your peers because we think you are a valuable source of information. If you assign the same rating to all persons in your group you are obviously not *carefully considering* each person as an *individual.* This makes your evaluation useless.

**Please make your decisions carefully.**

Enter your roster number here:


Enter your team number here:


DIRECTIONS:

You will be rating each member of your team (including yourself) in 5 specific areas: tactical performance, leadership, social interactions, physical fitness, and effort.

At the top of each page there are descriptions of what you should consider when making your ratings in each area. Read the information in the box carefully, then rate each member of your team in that area . When there are two ratings to make on a page, rate each person in the left column before continuing on to the right column.

This form will be read by a machine, so be sure to fill in your answers carefully and completely.

## EFFORT & PERSISTENCE

**LOW EFFORT Indicators:**

When he had trouble doing or learning something, he would give up.

He would get angry and frustrated if something didn't come easily to him.

He did not accept negative feedback well, and did not use it to improve his performance.

He was more concerned with just getting by or looking good than really learning something or doing his share of the work.

**HIGH EFFORT Indicators:**

When he was having trouble doing or learning something he would keep working at it.

If something didn't come easily to him, he would try even harder to learn it.

He accepted constructive feedback about his techniques or performance and used these to improve.

He would face any challenge with determination and never give up.

### LEVEL OF EFFORT/PERSISTENCE

| Roster Number | Low | Average | High |
|---|---|---|---|
| 1. _____ | | | |
| 2. _____ | | | |
| 3. _____ | | | |
| 4. _____ | | | |
| 5. _____ | | | |
| 6. _____ | | | |
| 7. _____ | | | |
| 8. _____ | | | |
| 9. _____ | | | |
| 10. _____ | | | |
| 11. _____ | | | |
| 12. _____ | | | |
| 13. _____ | | | |

Please list additional comments in this box:

TACTICAL PERFORMANCE

Factors to Consider

Reading contour/terrain maps
Finding land navigation points
Battle drills: reacting to contact, to an ambush, to indirect fire
Reconnaissance, combat, tracking patrols
Use of M60, AT4, Claymore mines

LEADERSHIP PERFORMANCE

Factors to Consider

Planning patrols
Issuing OPORDS
Directing, controlling, and supervising other team members
Focusing his squad on the task at hand
Coordinating squad actions
Using information and feedback from other team members

| Roster Number | TACTICAL PERFORMANCE Low | Average | High | LEADERSHIP PERFORMANCE Low | Average | High |
|---|---|---|---|---|---|---|
| 1. _____ | | | | | | |
| 2. _____ | | | | | | |
| 3. _____ | | | | | | |
| 4. _____ | | | | | | |
| 5. _____ | | | | | | |
| 6. _____ | | | | | | |
| 7. _____ | | | | | | |
| 8. _____ | | | | | | |
| 9. _____ | | | | | | |
| 10. _____ | | | | | | |
| 11. _____ | | | | | | |
| 12. _____ | | | | | | |
| 13. _____ | | | | | | |

Please list additional comments in this box:

3

Form Number 75020 5 72

SURVEY NETWORK™

SOCIAL INTERACTIONS

Factors to Consider:

How well he interacts with other team members regardless of rank or experience

Whether he provokes conflicts with other team members or works them out

Can he read people and situations well and act appropriately?

Does he know when to back off or be aggressive, when to be serious or funny?

PHYSICAL PERFORMANCE

Factors to Consider:

Strength
Endurance
Coordination
Ability to function with little sleep

| | SOCIAL INTERACTIONS | | | PHYSICAL PERFORMANCE | | |
|---|---|---|---|---|---|---|
| Roster Number | Low | Average | High | Low | Average | High |
| 1. _____ | | | | | | |
| 2. _____ | | | | | | |
| 3. _____ | | | | | | |
| 4. _____ | | | | | | |
| 5. _____ | | | | | | |
| 6. _____ | | | | | | |
| 7. _____ | | | | | | |
| 8. _____ | | | | | | |
| 9. _____ | | | | | | |
| 10. _____ | | | | | | |
| 11. _____ | | | | | | |
| 12. _____ | | | | | | |
| 13. _____ | | | | | | |

Please list additional comments in this box:

4

| Roster Number | Overall, how well did this soldier perform in Phase I? | | | After a year on an SF team, how well would you expect him to perform? | | |
|---|---|---|---|---|---|---|
| | PERFORMANCE IN PHASE I | | | FUTURE PERFORMANCE | | |
| | Poor | Good | Outstanding | Poor | Good | Outstanding |
| 1. _____ | | | | | | |
| 2. _____ | | | | | | |
| 3. _____ | | | | | | |
| 4. _____ | | | | | | |
| 5. _____ | | | | | | |
| 6. _____ | | | | | | |
| 7. _____ | | | | | | |
| 8. _____ | | | | | | |
| 9. _____ | | | | | | |
| 10. _____ | | | | | | |
| 11. _____ | | | | | | |
| 12. _____ | | | | | | |
| 13. _____ | | | | | | |

Please list additional comments in this box:

Form Number 75020-5-72 SURVEY NETWORK™

TEAM MEMBER EVALUATIONS

We understand that you are extremely tired, but we want your evaluation of the performance of each of your students in several areas. You may not have had the opportunity to completely evaluate each student in each of these areas, but we want your best estimate given the information you have.

Please make your decisions as carefully and accurately as possible. Information from this study may have implications for the way assessments are made in the future.

CADRE NAME:

_____

Enter your team's number here:


DIRECTIONS:

You will be rating each candidate on this team in 5 specific areas: tactical performance, leadership, social interactions, physical performance, and effort.

At the top of each page there are descriptions of what you should consider when making your ratings in each area. Read the information in the box carefully, then rate each student in that area. When there are two ratings to make on a page, rate each student in the left column before continuing on to the right column.

This form will be read by a machine, so be sure to fill in your answers carefully and completely.

# APPENDIX E

## Responsive Environment

DIRECTIONS:

The 10 statements below describe opinions a student might have about his experiences in this course. Indicate whether you agree or disagree with each statement by filling in the circle for the answer that is most appropriate for you.

Please answer honestly and remember that **no personnel** here at Camp Mackall or any part of the Special Warfare Center will ever see your individual responses or the responses of your team as a group. They will only see response averages for the class **as a whole**. Your responses to these questions can provide useful feedback about your experience here.

| | STRONGLY DISAGREE | SOMEWHAT DISAGREE | NEUTRAL | SOMEWHAT AGREE | STRONGLY AGREE |
|---|---|---|---|---|---|
| 1. There is not enough time for most students to learn and practice all of the things we are expected to know. | | | | | |
| 2. Most cadre here are very invested in the students' development. | | | | | |
| 3. There are things going on in my personal/family life that make it difficult to concentrate on training right now. | | | | | |
| 4. The students on my team tend to motivate each other. | | | | | |
| 5. There are not enough instructors to give the students all of the time and attention they need. | | | | | |
| 6. There are people on my team who really know what they are doing and are willing to help other people out. | | | | | |
| 7. The standards they use in this course seem fair and consistent. | | | | | |
| 8. The feedback and guidance students receive from instructors is constructive and helpful. | | | | | |
| 9. Most of the people on my team are only interested in themselves and looking good for the cadre. | | | | | |
| 10. My team generally pulls together to help each other out. | | | | | |

Please add any additional comments in this space.